



# K SCHOOL

## 20 Master de Data Science

Tableau y visualización – 4 y 5 de septiembre

Txema Fernández



# ¿Quién soy?



**Txema Fernández**

Big Data & Data Engineer

EL ARTE DE MEDIR



@txemaskapao

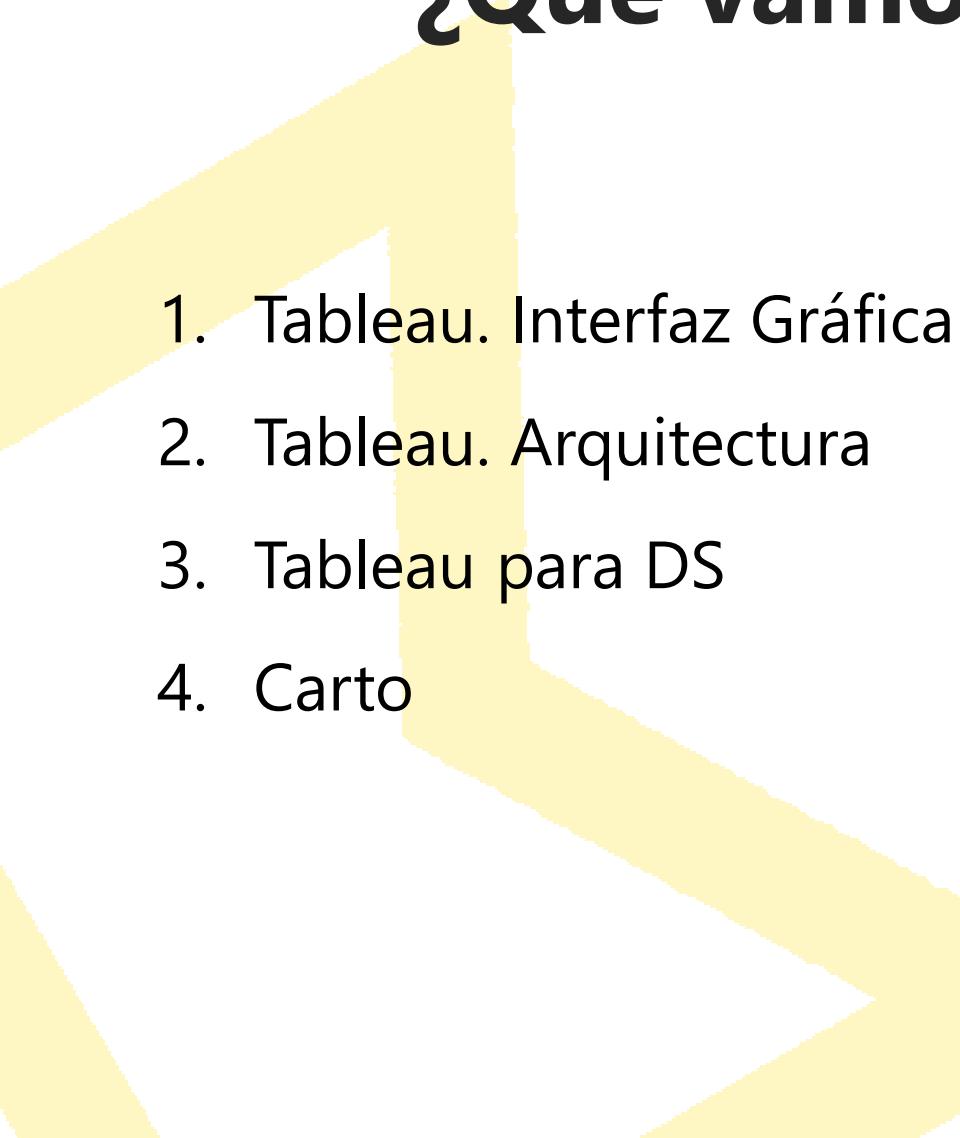


txema@elartedemedir.com



<https://www.linkedin.com/in/jmfdezfiano/>

# ¿Qué vamos a ver hoy?

- 
- 1. Tableau. Interfaz Gráfica
  - 2. Tableau. Arquitectura
  - 3. Tableau para DS
  - 4. Carto
- 1. Introducción
  - 2. Conexión a los datos
  - 3. Hojas de Trabajo
  - 4. Gráficos predeterminados
  - 5. Cálculos
  - 6. Detección de valores atípicos
  - 7. Combinación de datos
  - 8. Eje Doble
  - 9. Dashboards

# ¿Qué vamos a ver hoy?

- 1. Tableau. Interfaz Gráfica
- 2. Tableau. Arquitectura
- 3. Tableau para DS
- 4. Carto

- 1. Introducción
- 2. Conexión a los datos
- 3. Hojas de Trabajo
- 4. Gráficos predeterminados
- 5. Cálculos
- 6. Detección de valores atípicos
- 7. Combinación de datos
- 8. Eje Doble
- 9. Dashboards

## ¿Por qué Tableau?

- Es una herramienta de visualización de datos **extremadamente potente**, que permite llevar a cabo visualizaciones en cuestión de segundos.
- Su motor interno, basado en buenas prácticas, sugiere qué visualización se ajusta mejor a las métricas y dimensiones introducidas.
- Diseñar Dashboards es muy sencillo.

# PowerBi vs Tableau

FORRESTER® RESEARCH

### THE FORRESTER WAVE™

Enterprise BI Platforms (Vendor-Managed)

Q3 2019

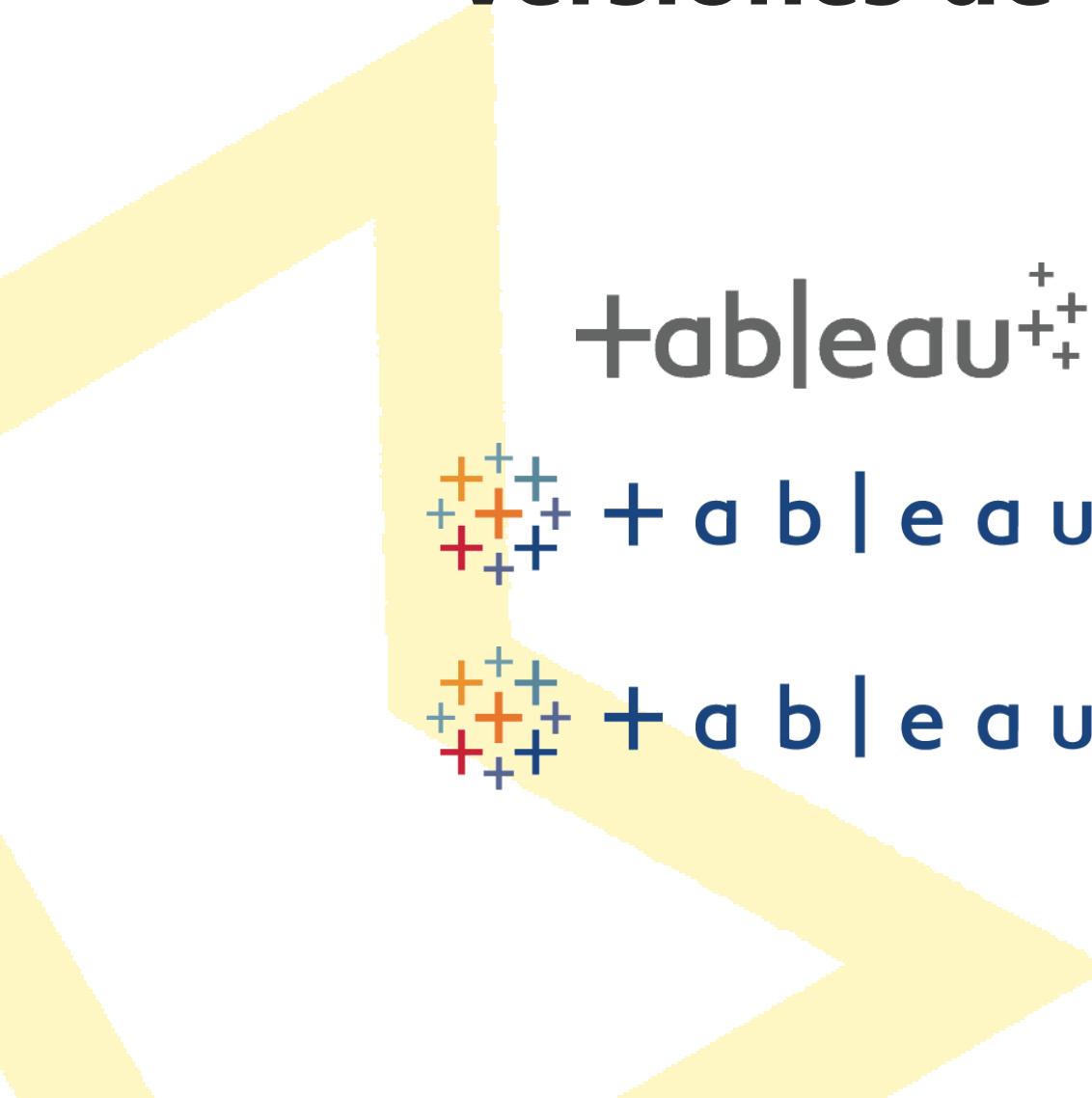


# PowerBi vs Tableau III

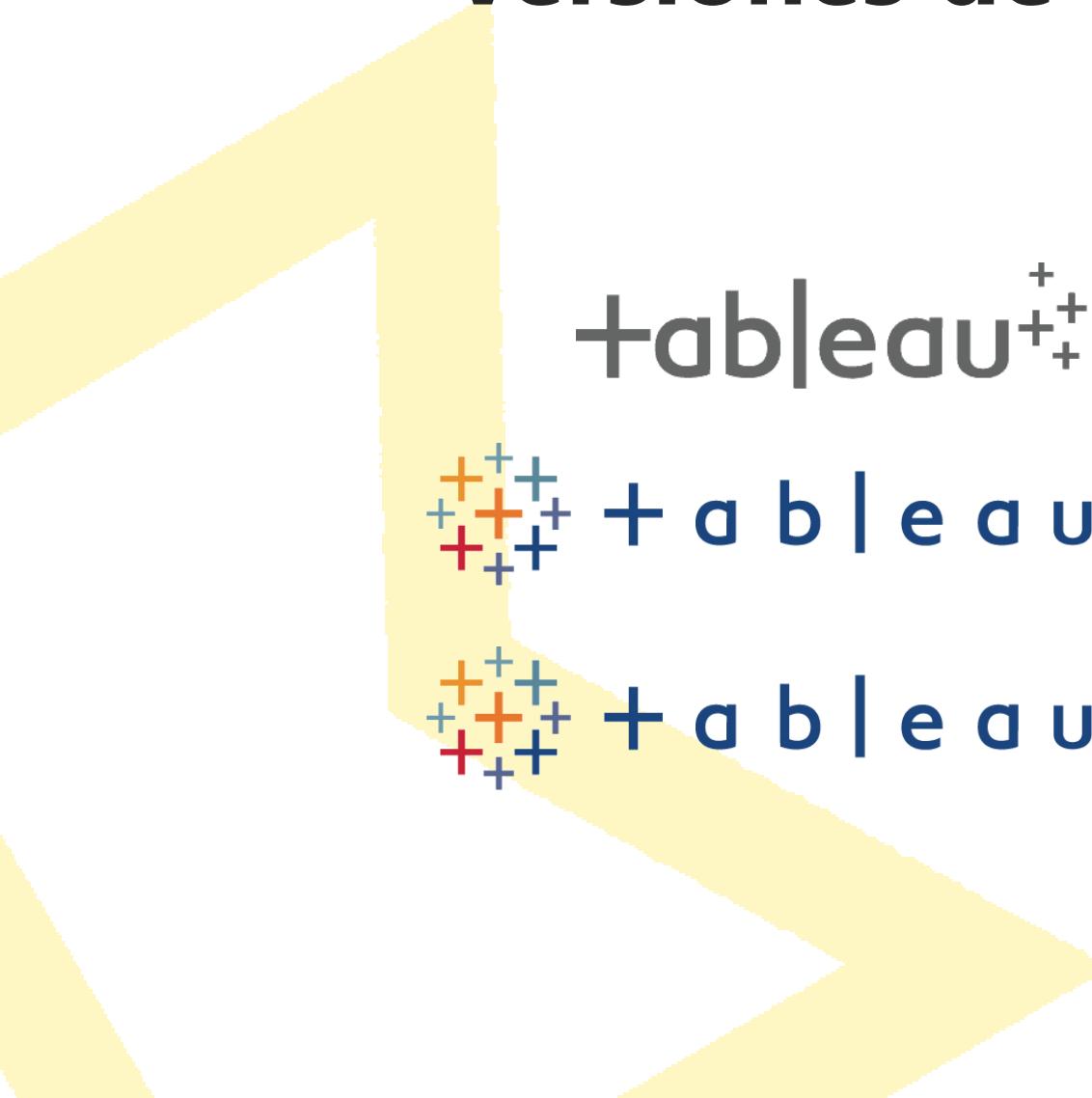


# Versiones de Tableau

tableau<sup>+</sup><sup>+</sup>public

A large, semi-transparent yellow arrow shape points from left to right across the slide.

+ tableau® Desktop

A large, semi-transparent yellow arrow shape points from left to right across the slide.

+ tableau® Server

## Versiones de Tableau

tableau<sup>+</sup><sup>+</sup>public

-  Gratuita.
-  Limitación: el set de datos debe tener menos de 1 millón de filas.
-  Los ficheros se almacenan en su web y no en tu PC.
-  Limitación: no se puede conectar a bases de datos. Solamente a Excel, archivos de texto y archivos de Microsoft Access.

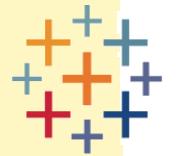
## Versiones de Tableau



+ a b | e a u® Desktop

-  Dos versiones: personal y profesional.
-  Ambas permiten guardar los archivos en tu PC.
-  Versión **personal**: Puedes guardar los datos en Tableau Online y Tableau Server.
-  Versión **profesional**: Única versión que se puede conectar a Bases de Datos y a Tableau Server.

## Versiones de Tableau



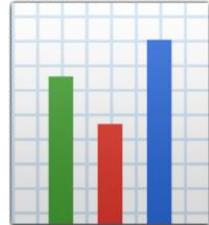
+ a b | e a u® Server

- Allows sharing of Dashboards, visualizations etc. within the company, so that people with permissions can access, collaborate, edit etc.

## Pero antes de comenzar...

### ¿Qué objetivo perseguimos?

Nuestra visualización debe ser capaz,  
por sí sola, de dar respuesta a la  
pregunta que nos formulemos

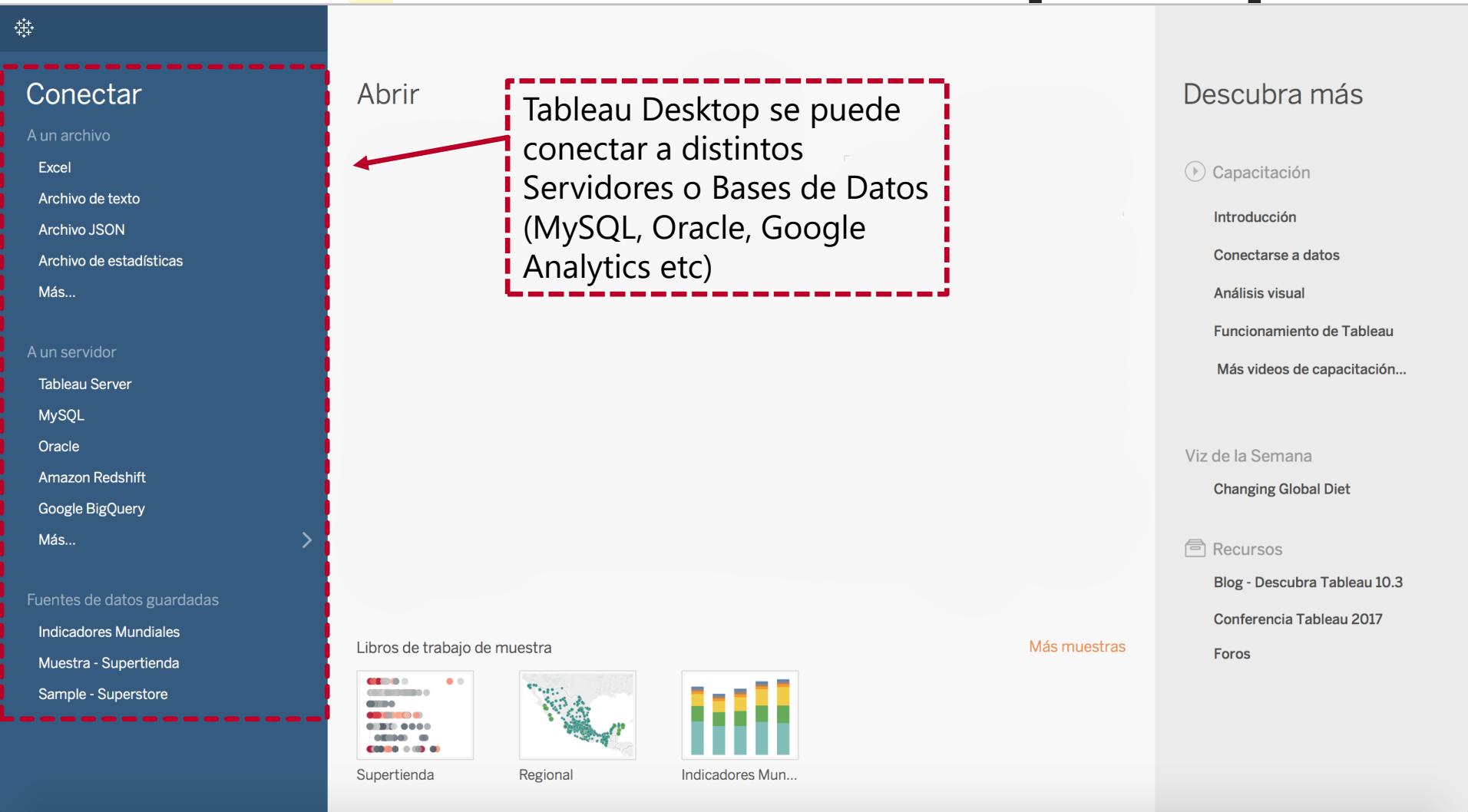


# ¿Qué vamos a ver hoy?

1. Tableau. Interfaz Gráfica
2. Tableau. Arquitectura
3. Tableau para DS
4. Carto

1. Introducción
2. Conexión a los datos
3. Hojas de Trabajo
4. Gráficos predeterminados
5. Cálculos
6. Detección de valores atípicos
7. Combinación de datos
8. Eje Doble
9. Dashboards

# Comencemos. Interfaz principal



Abrir

Tableau Desktop se puede conectar a distintos Servidores o Bases de Datos (MySQL, Oracle, Google Analytics etc)

Conectar

- A un archivo
- Excel
- Archivo de texto
- Archivo JSON
- Archivo de estadísticas
- Más...

A un servidor

- Tableau Server
- MySQL
- Oracle
- Amazon Redshift
- Google BigQuery
- Más...

Fuentes de datos guardadas

- Indicadores Mundiales
- Muestra - Supertienda
- Sample - Superstore

Libros de trabajo de muestra

- Supertienda
- Regional
- Indicadores Mun...

Más muestras

Descubra más

- Capacitación
- Introducción
- Conectarse a datos
- Análisis visual
- Funcionamiento de Tableau
- Más videos de capacitación...

Viz de la Semana

- Changing Global Diet

Recursos

- Blog - Descubra Tableau 10.3
- Conferencia Tableau 2017
- Foros

# Nuestra fuente de datos

The screenshot shows the Tableau Data Source interface. On the left, under 'Conexiones', there is a single entry: 'madrid-precio-distrito' (Excel). Under 'Hojas', there is also one entry: 'madrid-distritos'. A blue arrow points from the text 'Tableau automáticamente reconoce el tipo de variable. Aunque nos fiaremos la mayoría de las veces, Tableau tampoco es perfecto 😊' to the 'madrid-distritos' connection entry. Another blue arrow points from the text 'Estos son los nombres que Tableau ha interpretado automáticamente de nuestro dataset. Es la cabecera de nuestro fichero Excel' to the column headers in the preview area.

Conexiones Añadir

madrid-precio-distrito Excel

Hojas

Usar el intérprete de datos  
Puede que el intérprete de datos esté disponible para limpiar su libro de trabajo Excel.

madrid-distritos

Nueva unión

Conexión  En vivo  Extraer

Filtros 0 | [Añadir](#)

madrid-distritos (madrid-precio-distrito)

Tableau automáticamente reconoce el tipo de variable. Aunque nos fiaremos la mayoría de las veces, Tableau tampoco es perfecto 😊

Estos son los nombres que Tableau ha interpretado automáticamente de nuestro dataset. Es la cabecera de nuestro fichero Excel

madrid-distritos

Ordenar campos Orden de fuente de datos ▾

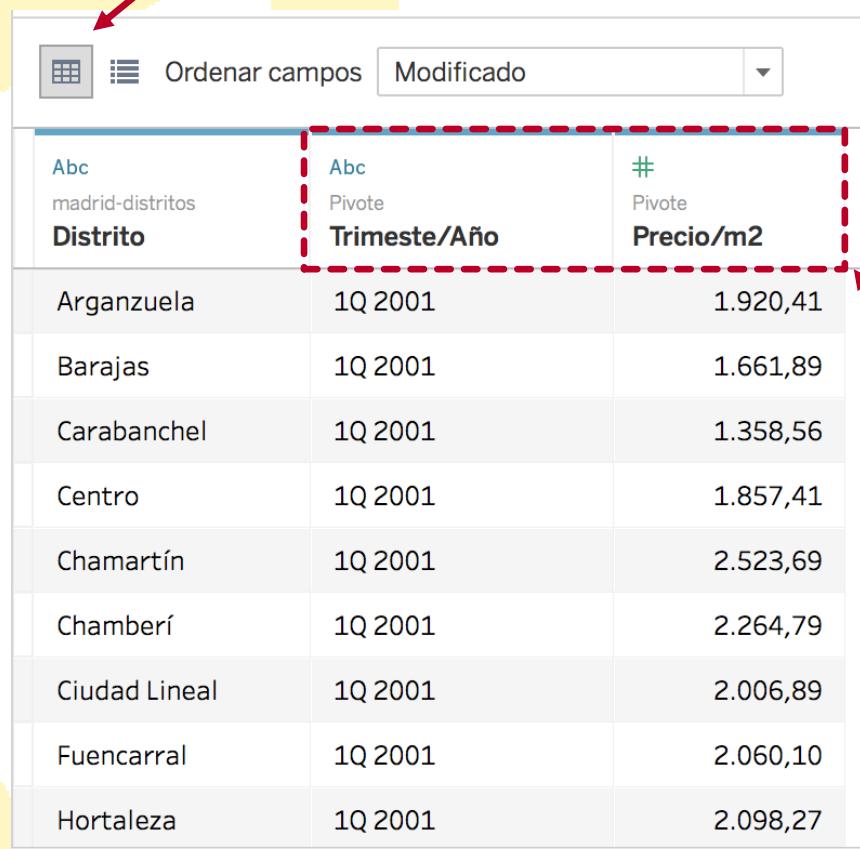
Mostrar alias Mostrar campos ocultos 21 filas

Distrito	4Q 2000	1Q 2001	2Q 2001	3Q 2001	4Q 2001	1Q 2002	2Q 2002	3Q 2002	4Q 2002
Arganzuela	n.d.	1.920,41	1.989,76	1.974,72	2.027,38	2.150,81	nulo	nulo	nulo
Barajas	n.d.	1.661,89	1.679,78	1.842,43	1.964,72	2.082,55	2.151,75	2.542,64	2.593,13
Carabanchel	n.d.	1.358,56	1.413,04	1.481,47	1.572,90	1.651,45	1.799,20	1.837,54	2.069,46
Centro	n.d.	1.857,41	1.988,01	2.061,03	2.155,49	2.292,02	2.454,73	2.523,00	2.789,52
Chamartín	n.d.	2.523,69	2.660,92	2.857,18	3.006,46	2.836,27	3.095,30	3.160,76	3.529,23
Chamberí	n.d.	2.264,79	2.454,96	2.446,56	2.695,06	2.780,16	2.901,73	3.180,40	3.353,40
Ciudad Lineal	n.d.	2.006,89	2.073,37	2.254,54	2.277,06	2.372,44	2.495,88	2.674,49	2.818,50
Ir a la hoja de trabajo	Quintana	n.d.	2.060,10	2.056,77	2.047,43	2.174,19	2.432,29	2.562,27	2.675,43
Montañesa	n.d.	2.098,27	2.148,78	2.104,17	2.192,90	2.179,80	2.445,38	2.422,00	2.757,72

Ir a la hoja de trabajo Fuente de datos Hoja 1

# Preparando nuestro Dataset

Pivatar la tabla: Nuestra **vista previa de la fuente de datos** ha cambiado



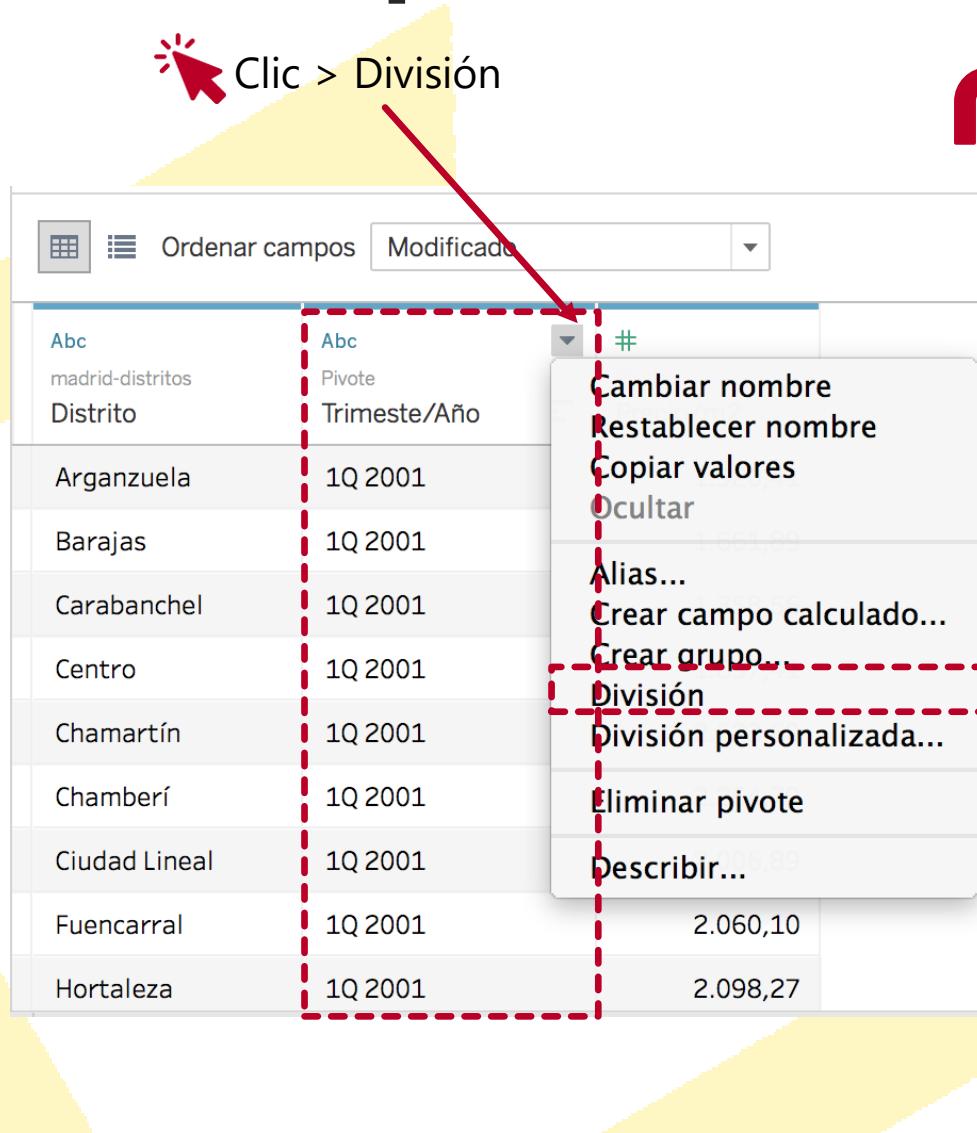
Abc madrid-distritos Distrito	Abc Pivot Trimestre/Año	# Pivot Precio/m2
Arganzuela	1Q 2001	1.920,41
Barajas	1Q 2001	1.661,89
Carabanchel	1Q 2001	1.358,56
Centro	1Q 2001	1.857,41
Chamartín	1Q 2001	2.523,69
Chamberí	1Q 2001	2.264,79
Ciudad Lineal	1Q 2001	2.006,89
Fuencarral	1Q 2001	2.060,10
Hortaleza	1Q 2001	2.098,27

Clic derecho > Cambiar nombre

Antes de comenzar, vamos a **cambiar el nombre** de las nuevas columnas que hemos pivotado, para que sean lo más explicativas posibles.

# Preparando nuestro Dataset

Clic > División



The screenshot shows a Tableau interface with a context menu open over a pivot field. The menu options include: Cambiar nombre, Restablecer nombre, Copiar valores, Ocultar, Alias..., Crear campo calculado..., Crear grupo..., División (highlighted with a red dashed box), División personalizada..., Eliminar pivotе, and Describir... . A yellow arrow points from the text "Clic > División" to the "División" menu item.

=Abc	=Abc
Cálculo	Cálculo
Trimestre	Año
1	2001
1	2001
1	2001
1	2001
1	2001
1	2001
1	2001
1	2001
1	2001
1	2001

De cara a facilitar el análisis, dividimos la variable Trimestre/Año en dos variables: trimestre y año

# Ejercicio



## ¿Qué vamos a ver?

- 
- A large, semi-transparent yellow chevron shape points from the bottom left towards the center of the slide, partially overlapping the list of topics.
- 1. Tableau. Interfaz Gráfica
  - 2. Tableau. Arquitectura
  - 3. Tableau para DS
  - 4. Carto
  - 1. Introducción
  - 2. Conexión a los datos
  - 3. Hojas de Trabajo
  - 4. Gráficos predeterminados
  - 5. Cálculos
  - 6. Detección de valores atípicos
  - 7. Combinación de datos
  - 8. Eje Doble
  - 9. Dashboards

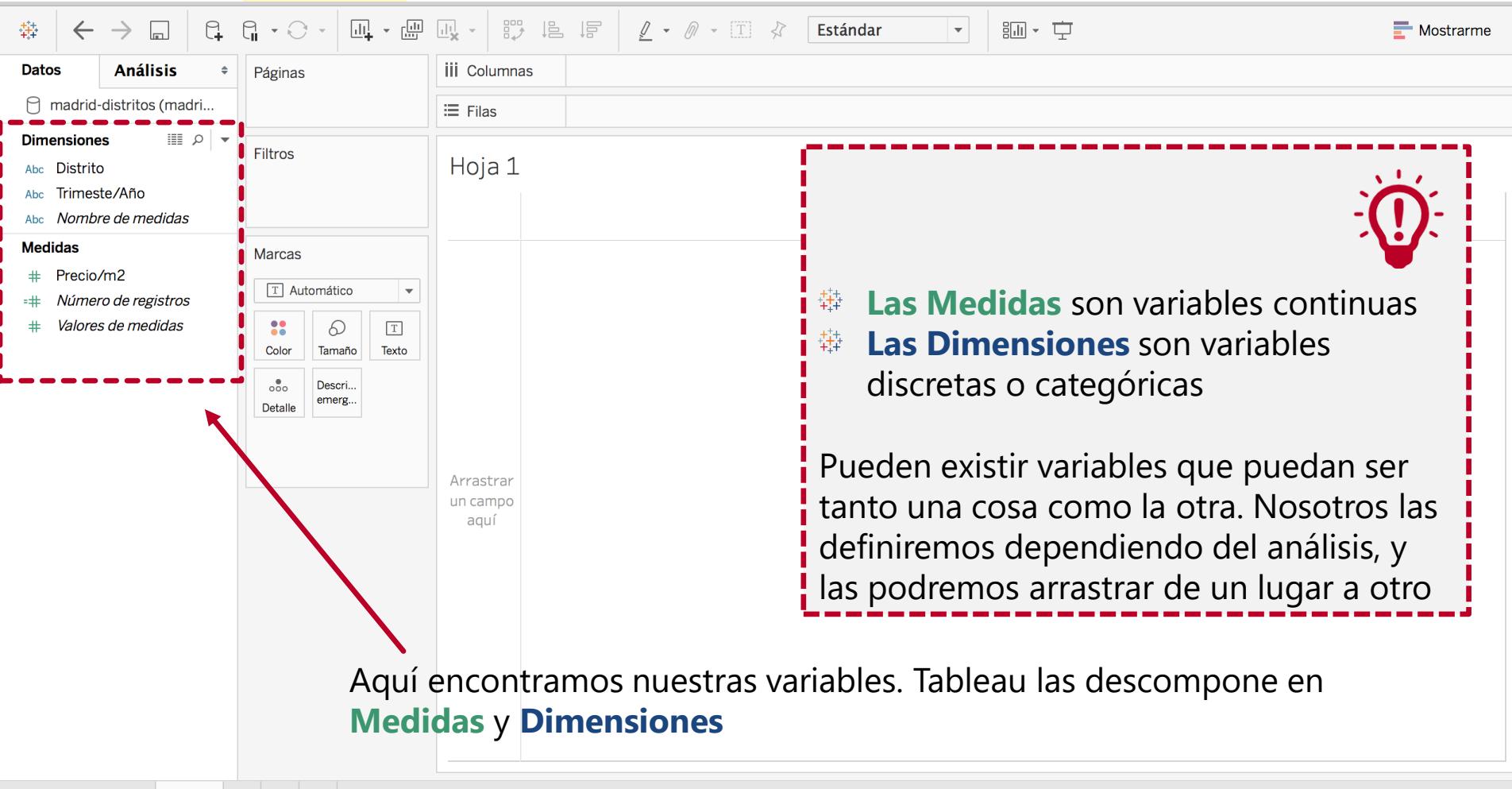
# Hojas de Trabajo



The screenshot shows the Tableau desktop application interface with several key components highlighted:

- Barra de herramientas (Toolbar):** Located at the top, it includes icons for file operations, zoom, and various analysis tools.
- Panel Mostrarme (Show Me Panel):** Located on the right side, it contains a large red box around the "Mostrar me" button.
- Variables (Variables):** A sidebar on the left containing sections for "Dimensiones" (Dimensions) and "Medidas" (Measures). Dimensions include "Distrito", "Trimestre/Año", and "Nombre de medidas". Measures include "Precio/m2", "Número de registros", and "Valores de medidas".
- Estantes (Shelves):** A central area divided into three shelves:
  - Páginas (Pages):** Shows "Hoja 1".
  - Columnas (Columns):** Contains a red box around the "Filas" (Rows) section.
  - Filtros (Filters):** Contains a red box around the "Marcas" (Marks) section, which includes options like "Automático", "Color", "Tamaño", "Texto", "Detalle", and "Descri... emerg...".
- Espacio de trabajo (Work Area):** The main workspace where "Arrastrar un campo aquí" (Drag a field here) is indicated multiple times.

# Medidas y dimensiones



Aquí encontramos nuestras variables. Tableau las descompone en **Medidas** y **Dimensiones**

**Las Medidas** son variables continuas  
**Las Dimensiones** son variables discretas o categóricas

Pueden existir variables que puedan ser tanto una cosa como la otra. Nosotros las definiremos dependiendo del análisis, y las podremos arrastrar de un lugar a otro

! Mostrarme

Datos Análisis Páginas Columnas Filas Estándar

madrid-distritos (madri...)

Dimensiones Distrito Trimestre/Año Nombre de medidas

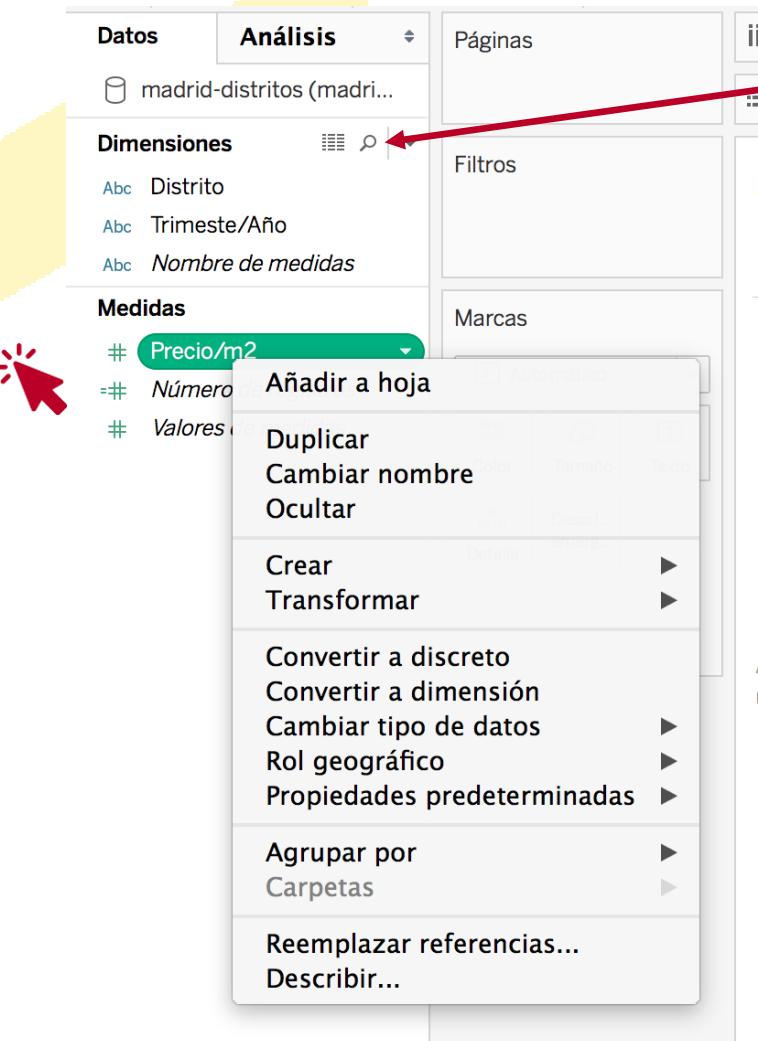
Medidas Precio/m<sup>2</sup> Número de registros Valores de medidas

Filtros Automático Color Tamaño Texto Detalle Descri... emerg...

Hoja 1

Arrastrar un campo aquí

# Medidas y dimensiones



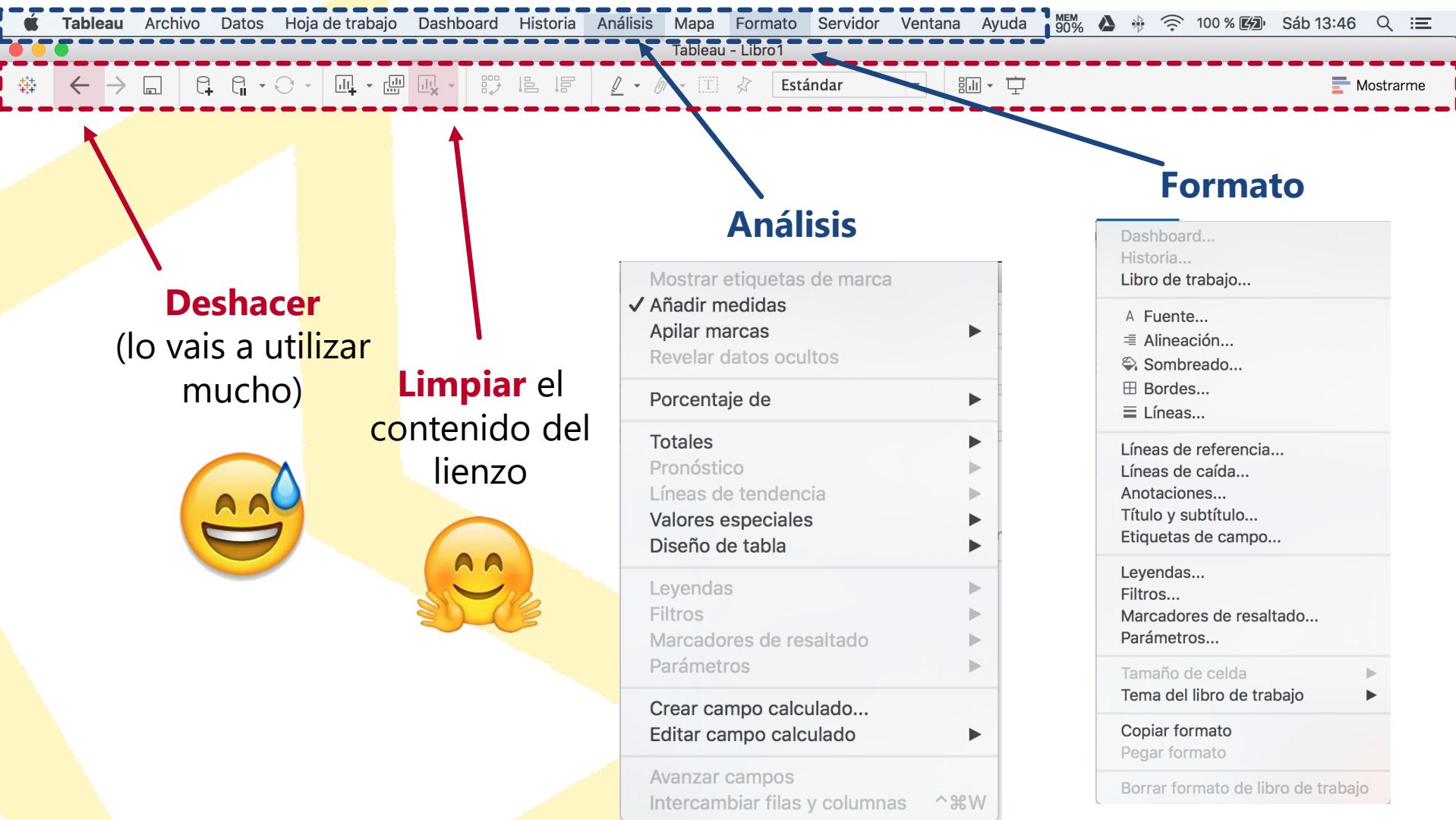
The screenshot shows the Tableau interface with the 'Análisis' tab selected. On the left, the 'Dimensiones' shelf lists 'Distrito', 'Trimestre/Año', and 'Nombre de medidas'. The 'Medidas' shelf lists 'Precio/m2' (selected), 'Número', and 'Valores'. A red cursor icon is on the 'Medidas' shelf. A context menu is open over 'Precio/m2', listing options like 'Añadir a hoja', 'Duplicar', 'Cambiar nombre', 'Ocultar', etc. A red arrow points from the search icon in the top right of the interface to the 'Nombre de medidas' measure in the dimensions shelf.

Cuando trabajemos **con infinidad de variables**, podremos buscar la que nos interese de forma sencilla

Haciendo clic derecho podemos ver una serie de acciones que aplican tanto a métricas como a dimensiones.

Podemos desde cambiarles el nombre, crear grupos, jerarquizarlas, asignarles un alias etc. hasta cambiar el tipo de variable (numérica, texto, fecha) que queramos que sea

# Barra de Herramientas



The screenshot shows the Tableau application window. At the top is the ribbon menu with tabs: Tableau, Archivo, Datos, Hoja de trabajo, Dashboard, Historia, Análisis, Mapa, Formato, Servidor, Ventana, Ayuda. Below the ribbon is the title bar "Tableau - Libro1". To the right of the title bar are system status icons: MEM 90%, battery level, signal strength, 100%, Sáb 13:46, and a search icon. The main workspace contains several floating toolbars: "Deshacer" (Undo), "Limpiar el contenido del lienzo" (Clear canvas content), "Análisis" (Analysis), "Formato" (Format), and "Mostrarme" (Show me). The "Análisis" toolbar is expanded, showing options like "Mostrar etiquetas de marca", "Añadir medidas", "Apilar marcas", "Revelar datos ocultos", etc. The "Formato" toolbar is also expanded, listing categories such as Fuente, Alineación, Sombreado, Bordes, Líneas, Líneas de referencia, Líneas de caída, Anotaciones, Título y subtítulo, Etiquetas de campo, Leyendas, Filtros, Marcadores de resultado, Parámetros, and more.

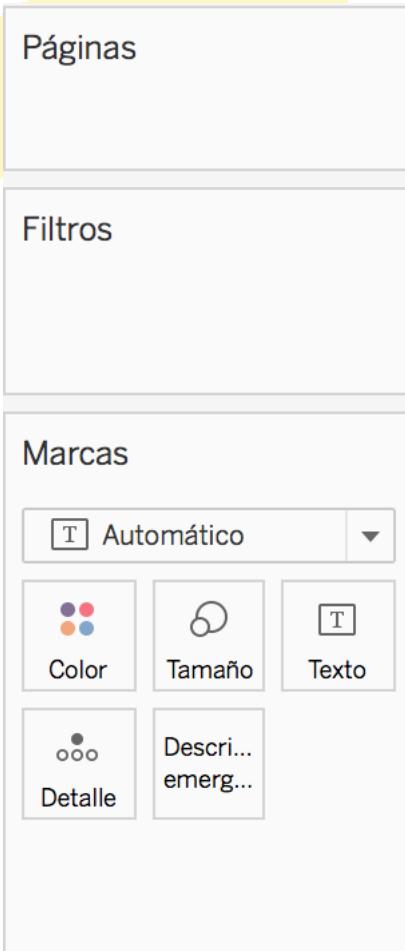
**Deshacer**  
(lo vais a utilizar  
mucho)

**Limpiar el**  
**contenido del**  
**lienzo**

**Análisis**

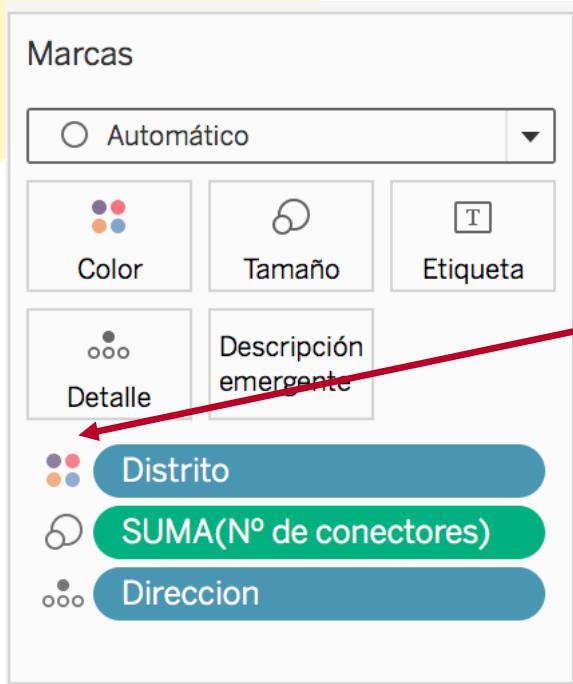
**Formato**

## Tarjetas



- **Páginas:** Permite dividir una vista en una serie de páginas para que pueda analizar mejor cómo un campo específico afecta al resto de los datos en una vista. Es lo más similar a contar una historia 😊
- **Filtros:** Permite especificar qué datos incluir y excluir. Por ejemplo, podría querer analizar el precio/m<sup>2</sup> para todos los distritos, pero solo para ciertos momentos temporales. Al colocar campos en el estante Filtros, puede crear tal vista.
- **Marcas:** Aquí arrastramos las distintas variables para controlar las propiedades de las marcas, como tipo, color, tamaño, forma, etc. Cada campo tiene un ícono junto a la tarjeta para identificar la propiedad de marca que está definiendo.

## Tarjetas | Marcas



Esta tarjeta es donde arrastramos nuestras distintas variables (Drag & Drop) para controlar propiedades de nuestra visualización.

A cada campo le acompaña un ícono que determina la propiedad que está definiendo a una marca.

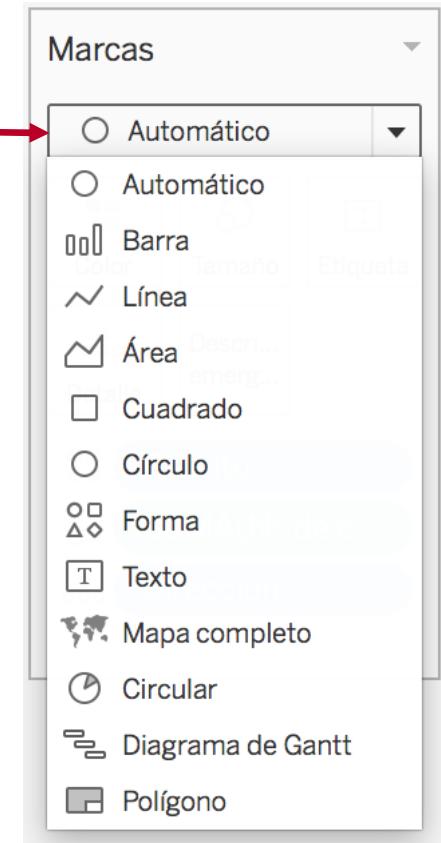
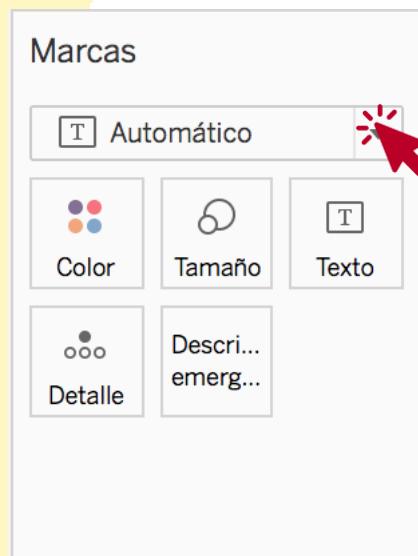
En nuestro ejemplo:

 Distrito -> Color

 N° de conectores -> Tamaño de las marcas

 Dirección -> Detalle

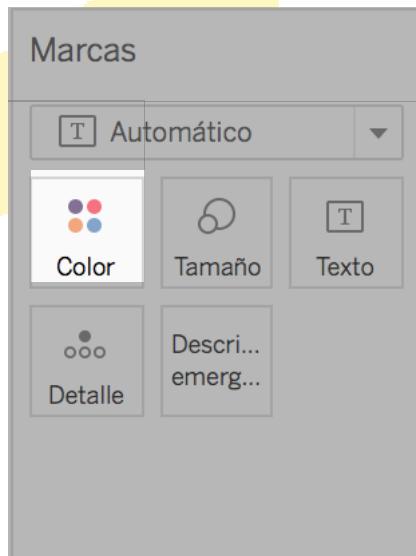
## Tarjetas | Marcas



Por defecto, el menú aparecerá en **Automático**. En ese caso, Tableau selecciona el tipo de marcas que mejor se ajusta en función a los datos introducidos en los estantes de Filas y Columnas.

Podemos “forzar” nuestra visualización para que nos muestre los datos de una u otra manera (barras, líneas, círculos, formas específicas etc)

# Tarjetas | Marcas | Color



💡 Por defecto, si no modificamos este campo, el azul será el color para todos nuestros gráficos y el negro para el texto.

💡 Cuando arrastramos un campo sobre la marca Color, Tableau aplica diferentes colores en función de los valores que tenga el campo. Pueden suceder dos cosas:

- 💡 Que arrastremos una **Medida**: Tableau muestra una escala de colores.
- 💡 Que arrastremos una **Dimensión**: Tableau asigna un color diferente para cada valor que tome ese campo.

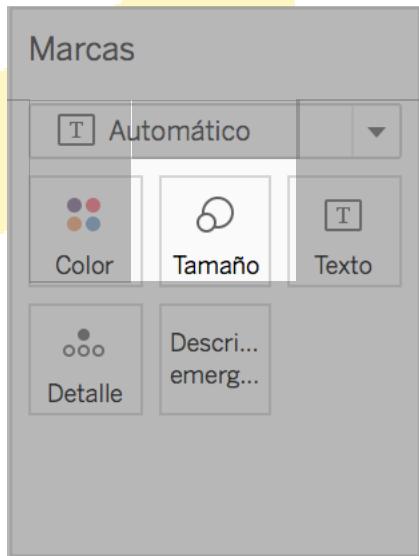
## Medida



## Dimensión

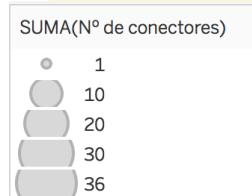
Distrito	?
Arganzuela	
Centro	
Chamartín	
Chamberí	

# Tarjetas | Marcas | Tamaño



¡Muy útil!

Medida



Dimensión

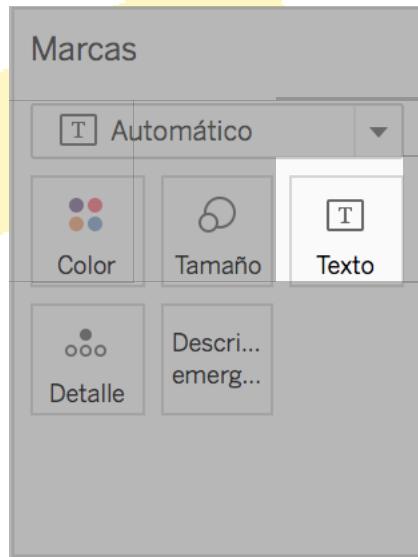


Nos permite asignar diferentes tamaños a las marcas de nuestras visualizaciones.

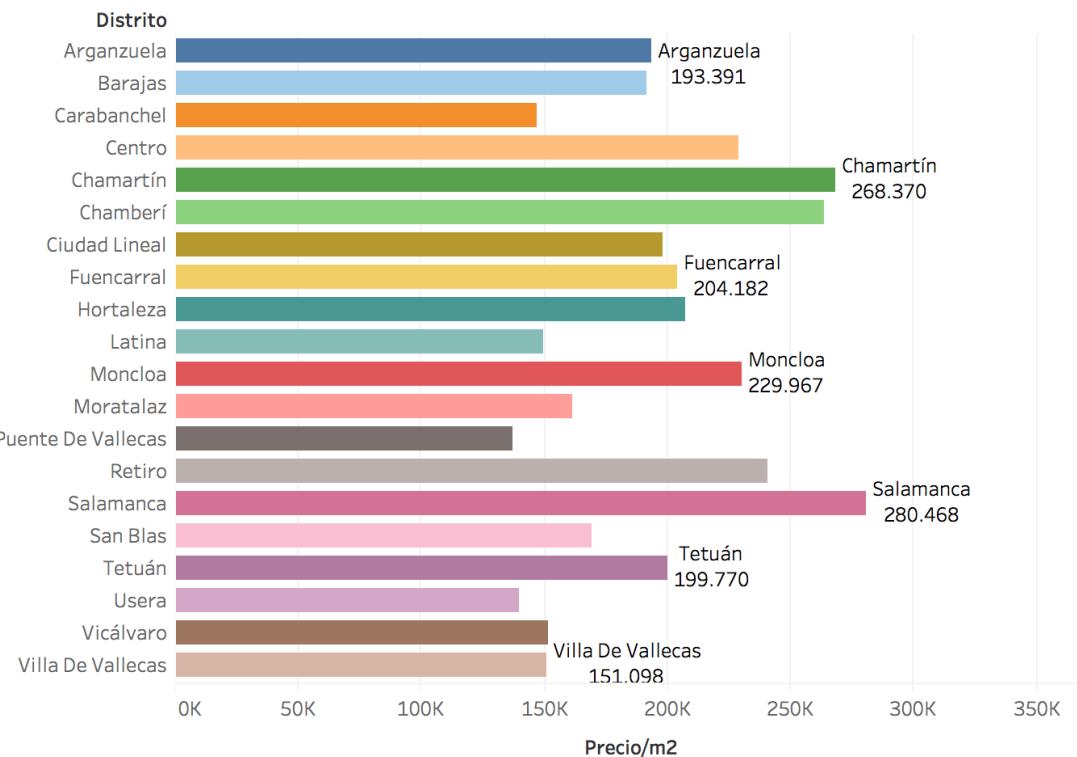
Pueden suceder dos cosas:

- Que arrastremos una **Medida**: Al ser un campo cuantitativo, cada marca pasará a adquirir un tamaño en función del campo asociado
- Que arrastremos una **Dimensión**: Tableau asigna un tamaño en función del número de miembros que haya. A cada miembro, un tamaño diferente.

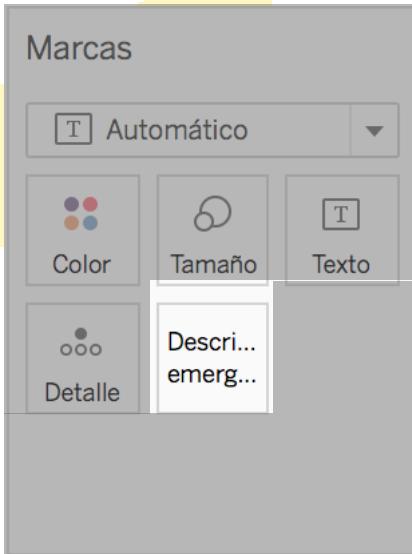
# Tarjetas | Marcas | Texto



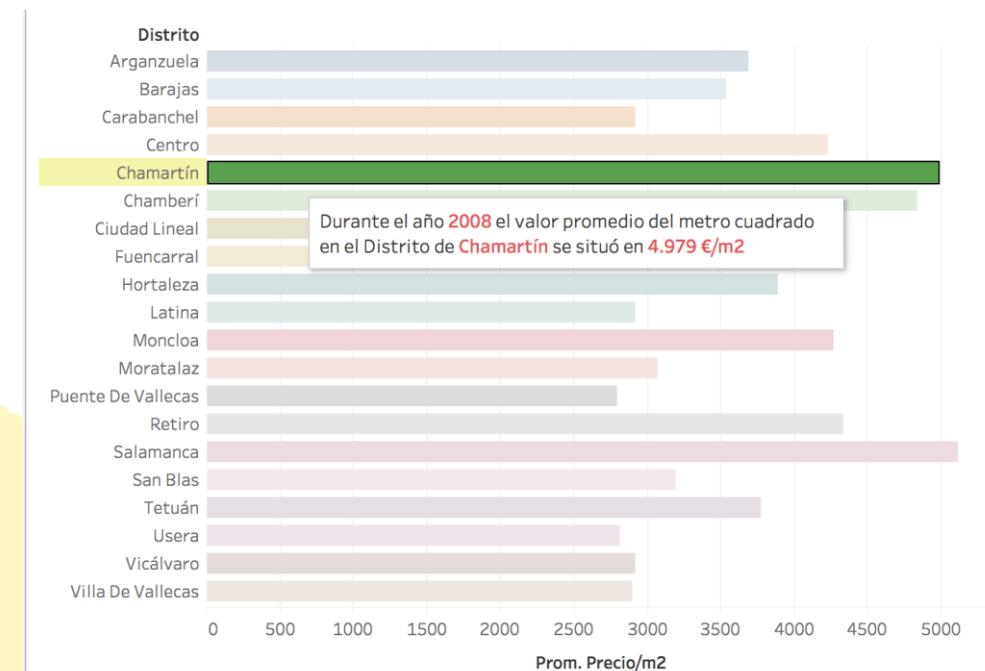
- Nos permite asignar un texto específico a las marcas de nuestras visualizaciones.
- Podemos agregar tanto **Medidas** como **Dimensiones**.



# Tarjetas | Marcas | Descrip. Emergente



- Nos ofrecen más información sobre los datos, al colocar el cursor encima de nuestras visualizaciones.
- Son especialmente útiles cuando utilizamos Dashboards dinámicos, ya que permiten detallar al máximo la información de modo que el receptor pueda entender perfectamente lo que tiene delante.



## Estantes

iii Columnas	Crea las columnas de una tabla
≡ Filas	Crea las filas de una tabla.

Podemos colocar todos los campos que queramos en estos estantes.



Al colocar una dimensión en los estantes Filas o Columnas, automáticamente se generan los encabezados de los miembros de dicha dimensión. Si por el contrario colocamos una medida en el estante Filas o Columnas, automáticamente se generan ejes cuantitativos para esa medida.

Cuantos más campos agreguemos a la vista, más encabezados y ejes adicionales se incluirán en la tabla, e iremos obteniendo una imagen cada vez más detallada de los datos.

## Panel Mostrarme



- ⊕ Nos muestra todas las distintas opciones y todas las distintas maneras en las que poder representar las variables que tenemos dentro de nuestro área de trabajo.
- ⊕ Automáticamente, Tableau nos dice el tipo de gráfico que se adapta a la combinación de medidas y dimensiones que hemos introducido en los estantes.
- ⊕ Cuando el tipo de gráfico aparece sombreado, no podemos utilizar ese tipo de visualización. Tableau quiere asegurarse que nuestros diseños siguen el manual de buenas prácticas



# Mostrarme | Tabla



Para tablas de texto pruebe

1 o más **Dimensiones**

1 o más **Medidas**

Distrito	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
	Año									
Arganzuela	1.978	2.151	2.933	3.300	3.568	3.873	3.926	3.688	3.398	3.322
Barajas	1.787	2.343	2.809	3.056	3.386	3.555	3.677	3.539	3.277	3.233
Carabanchel	1.456	1.839	2.313	2.615	2.774	3.102	3.165	2.917	2.629	2.466
Centro	2.015	2.515	3.061	3.467	3.833	4.248	4.420	4.230	3.918	3.777
Chamartín	2.762	3.155	3.787	4.158	4.389	4.694	4.897	4.979	4.789	4.696
Chamberí	2.465	3.054	3.678	4.132	4.372	4.745	4.940	4.839	4.693	4.621
Ciudad Line..	2.153	2.590	3.109	3.308	3.552	3.777	3.761	3.678	3.492	3.315
Fuencarral	2.085	2.619	3.064	3.370	3.510	3.818	3.890	3.709	3.540	3.443
Hortaleza	2.136	2.451	2.898	3.224	3.448	3.728	3.875	3.890	3.807	3.680
Latina	1.555	1.954	2.393	2.598	2.915	3.178	3.180	2.919	2.592	2.451
Moncloa	2.173	2.694	3.182	3.593	3.837	4.158	4.372	4.271	4.093	3.957
Moratalaz	1.739	2.072	2.558	2.848	3.019	3.323	3.322	3.069	2.750	2.609
Puente De ..	1.443	1.753	2.244	2.567	2.805	3.090	3.075	2.795	2.420	2.225
Retiro	2.356	2.790	3.346	3.729	3.963	4.274	4.415	4.331	4.170	4.218
Salamanca	2.603	3.189	3.832	4.207	4.506	4.739	5.024	5.116	4.966	4.802
San Blas	1.767	2.108	2.634	2.963	3.231	3.399	3.412	3.196	2.933	2.781
Tetuán	2.232	2.692	3.155	3.476	3.712	3.943	3.935	3.776	3.504	3.415
Usera	1.392	1.804	2.279	2.564	2.783	3.039	3.017	2.816	2.491	2.295
Vicálvaro	1.596	2.013	2.456	2.778	3.006	3.206	3.124	2.919	2.601	2.527
Villa De Vall..	1.395	1.691	2.193	2.425	2.719	3.079	3.123	2.900	2.632	2.500
Villaverde	1.228	1.551	1.938	2.405	2.575	2.777	2.826	2.579	2.304	2.125

+++ Son muy visuales.

+++ Podemos agregar multitud de variables, así como ordenar la tabla en función del campo que nos interese.

+++ Podemos aplicar otra serie de elementos, como color, para dar visibilidad y resaltar los datos

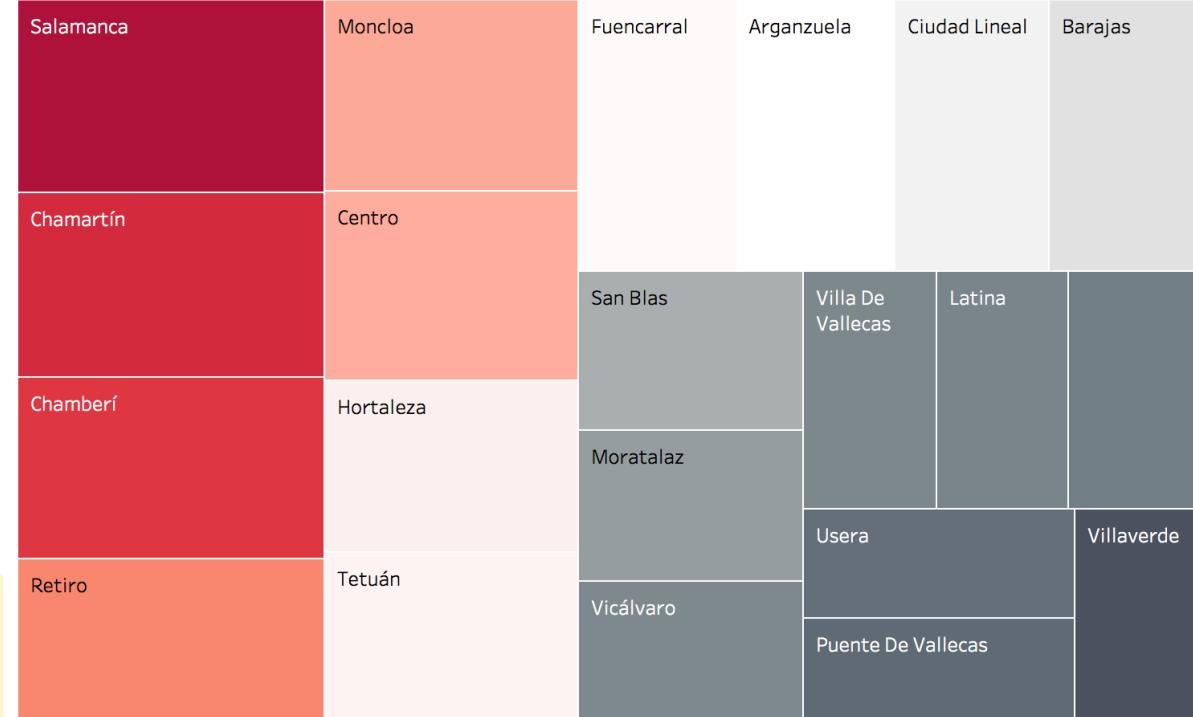
# Mostrarme | Mapa de árbol



Para tablas de texto pruebe

1 o más **Dimensiones**

1 o más **Medidas**



- ⊕⊕ Nos permite detectar de manera rápida y eficaz aquellos valores de mi variable más destacados.
- ⊕⊕ Permite asignar tintes cualitativos, gracias a la escala de color, a nuestras visualizaciones.

# Mostrarme | Burbujas agrupadas



Para tablas de texto pruebe

1 o más Dimensiones

1 o más Medidas

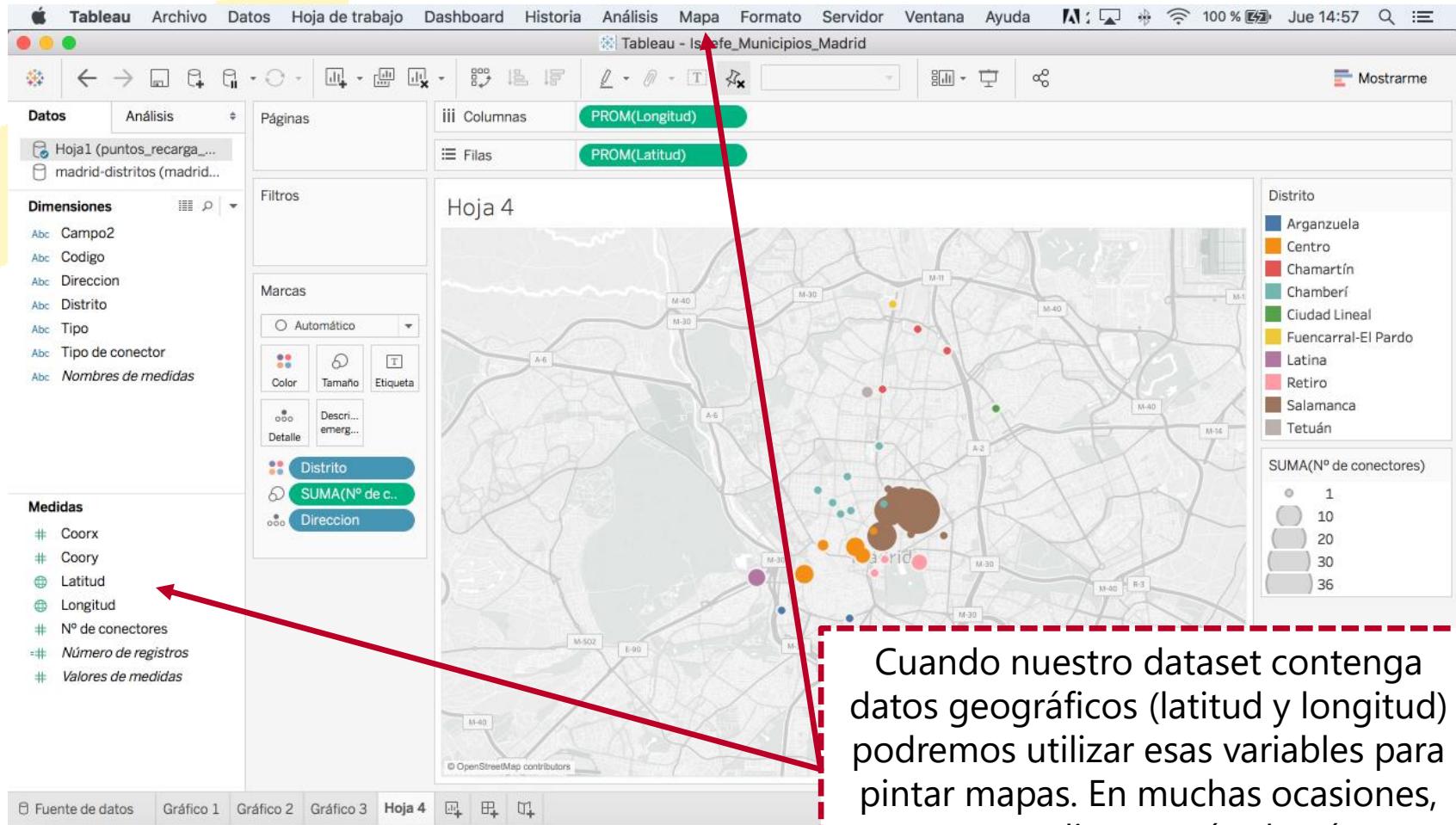


- Similar al mapa de árbol, ya que nos permite relacionar tanto colores como tamaños.
- Dependiendo del número de resultados que arroje la variable, estos gráficos serán más o menos difíciles de interpretar

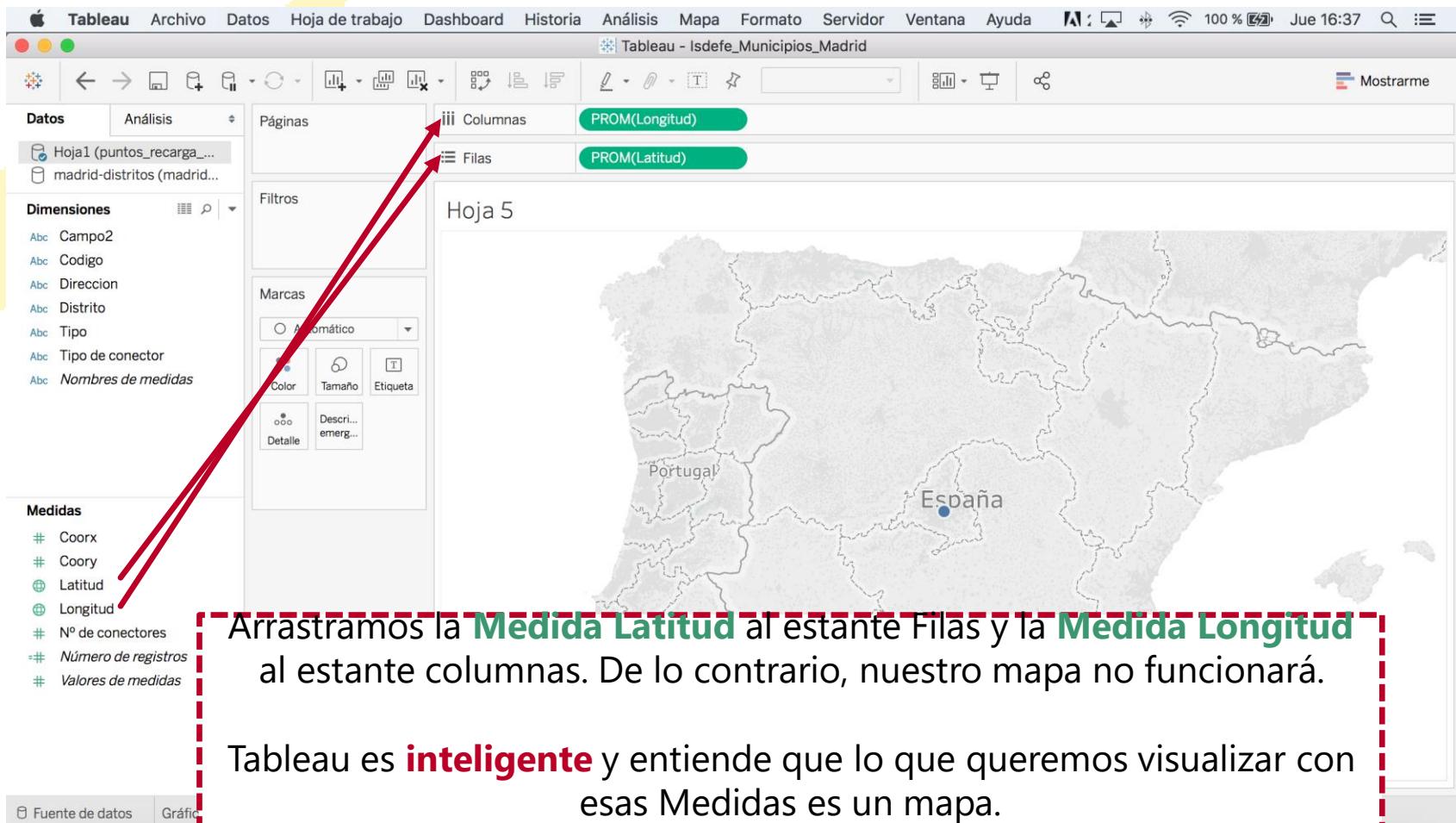
# Ejercicio



# Mapas



# Paso 1: Arrastrar nuestras medidas

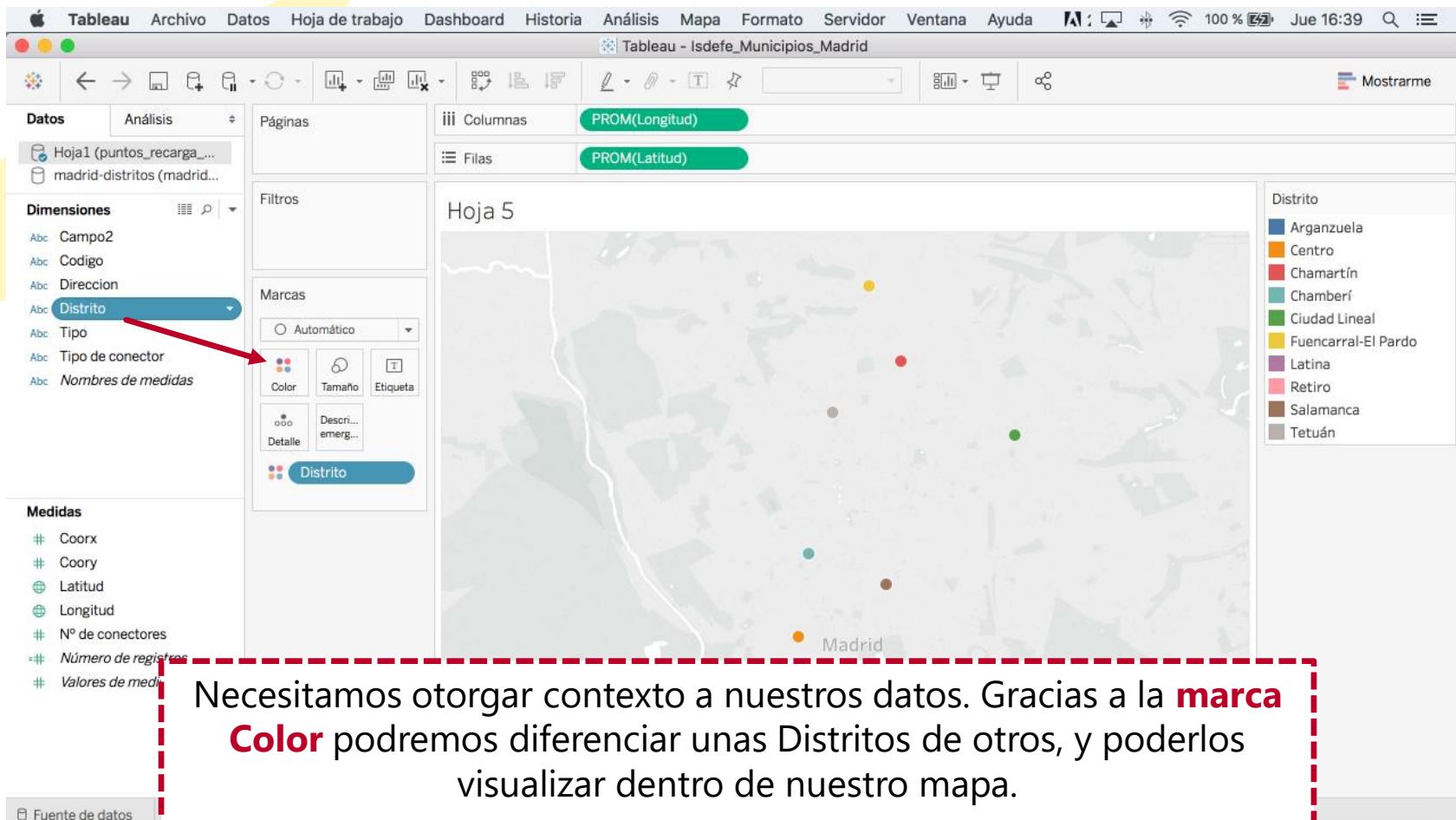


Arrastramos la **Medida Latitud** al estante Filas y la **Medida Longitud** al estante columnas. De lo contrario, nuestro mapa no funcionará.

Tableau es **inteligente** y entiende que lo que queremos visualizar con esas Medidas es un mapa.

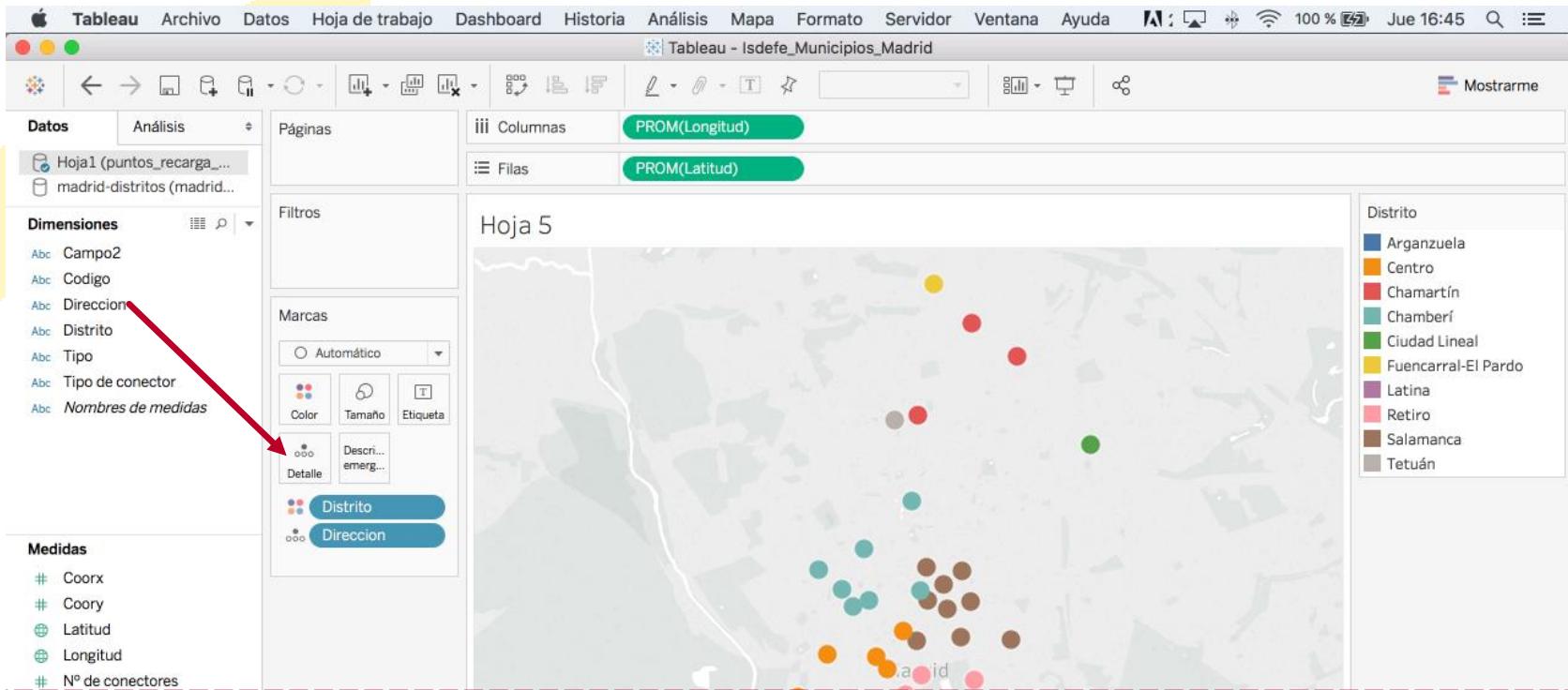
**¿Qué tenemos?** La latitud y longitud promedio de nuestro set de datos.

## Paso 2: Contextualizar



**¿Qué tenemos?** La latitud y longitud promedio de los distintos Distritos dentro de nuestro set de datos.

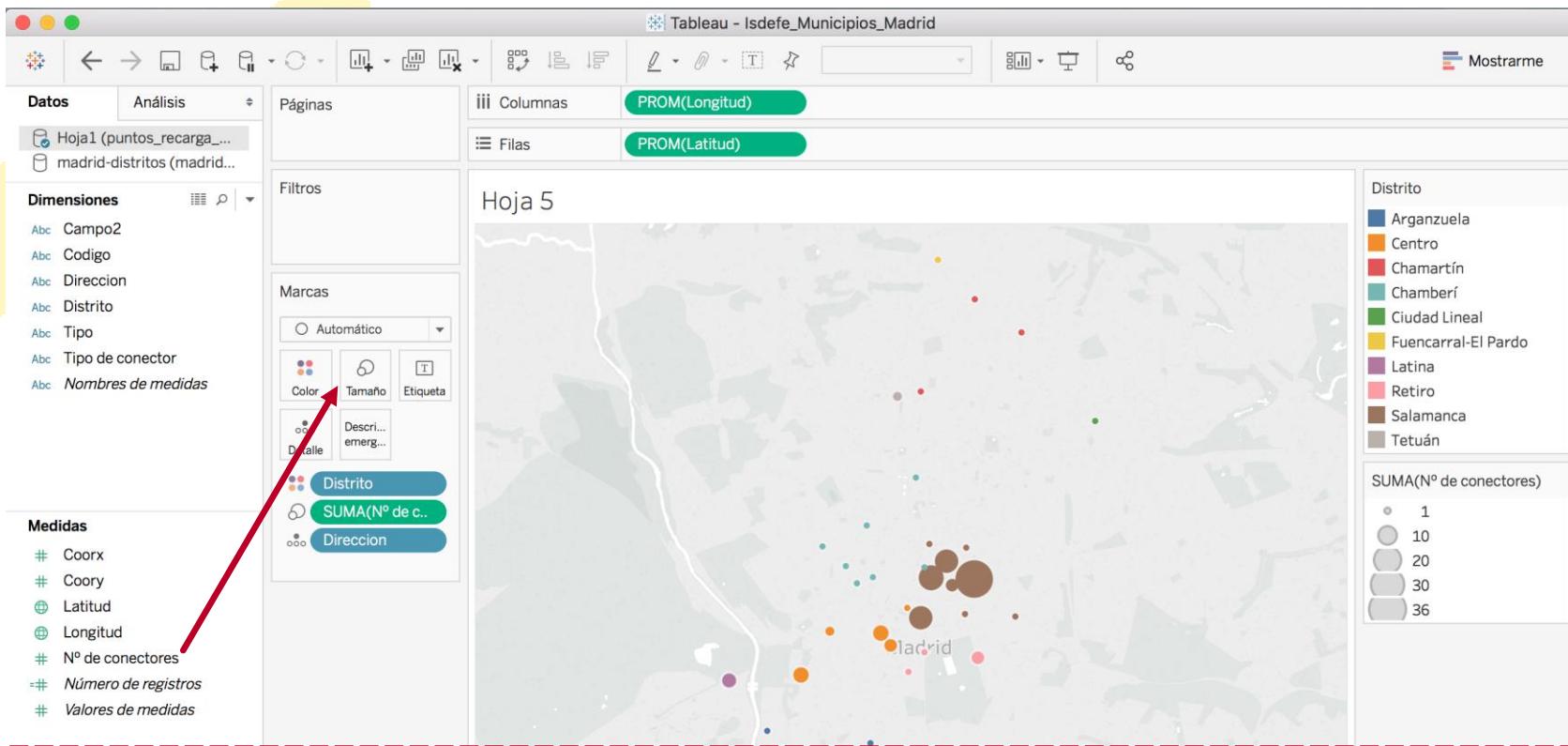
## Paso 3: Ir al detalle



Con la **marca Detalle** hemos conseguido no sólo los distritos dónde se encuentran las estaciones de recarga, sino también la dirección específica de cada una de ellas. Pero aún nos falta algo...

¿Dónde se encuentran las estaciones de recarga? ✓  
¿Cuántos puntos de recarga hay por estación? ✗

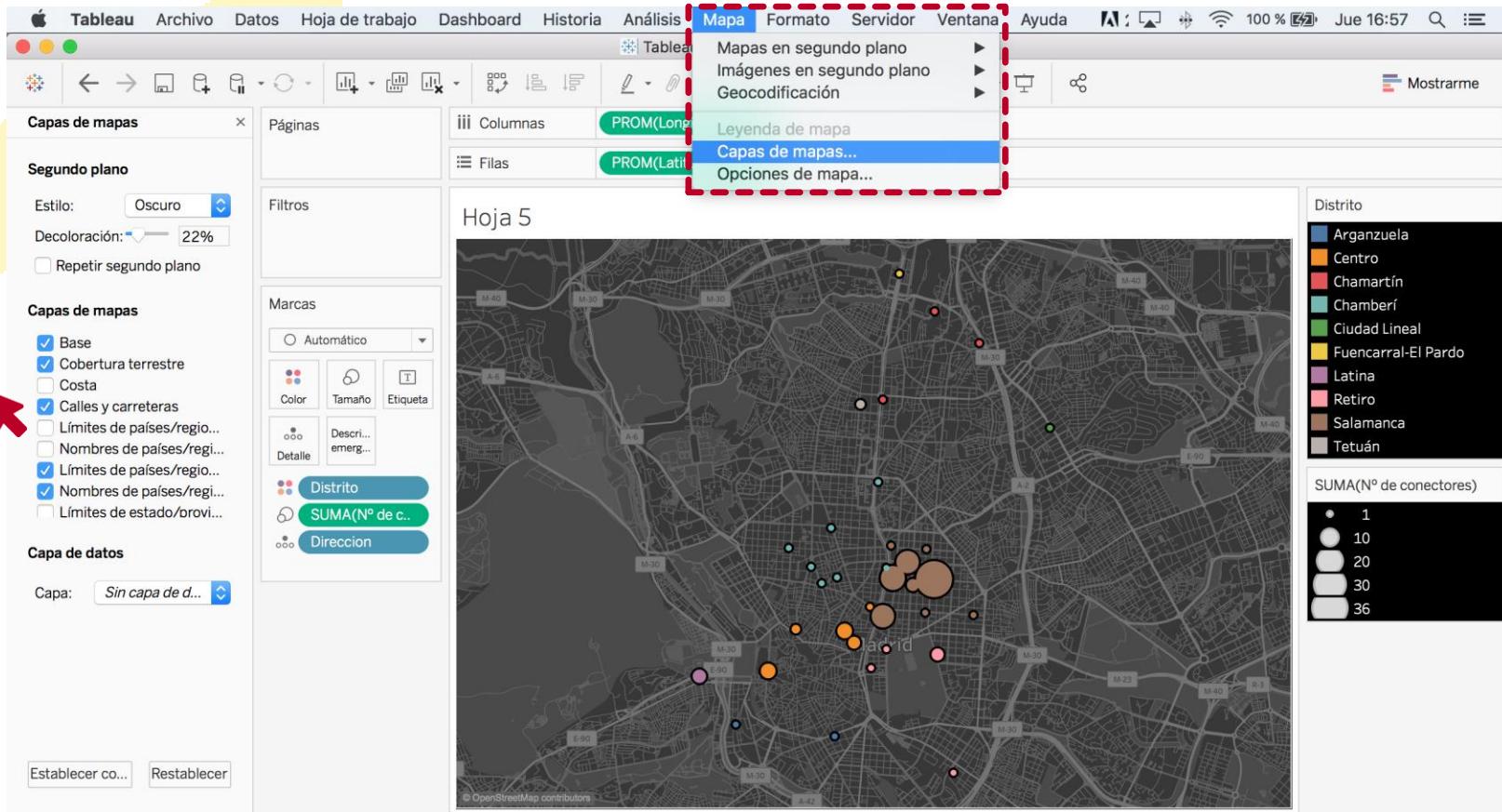
# Paso 4: Responder a la pregunta



Agregando el Nº de conectores a la **marca Tamaño** pasamos a conocer, de forma visual, no sólo dónde están las estaciones sino cuántos puntos de recarga hay.

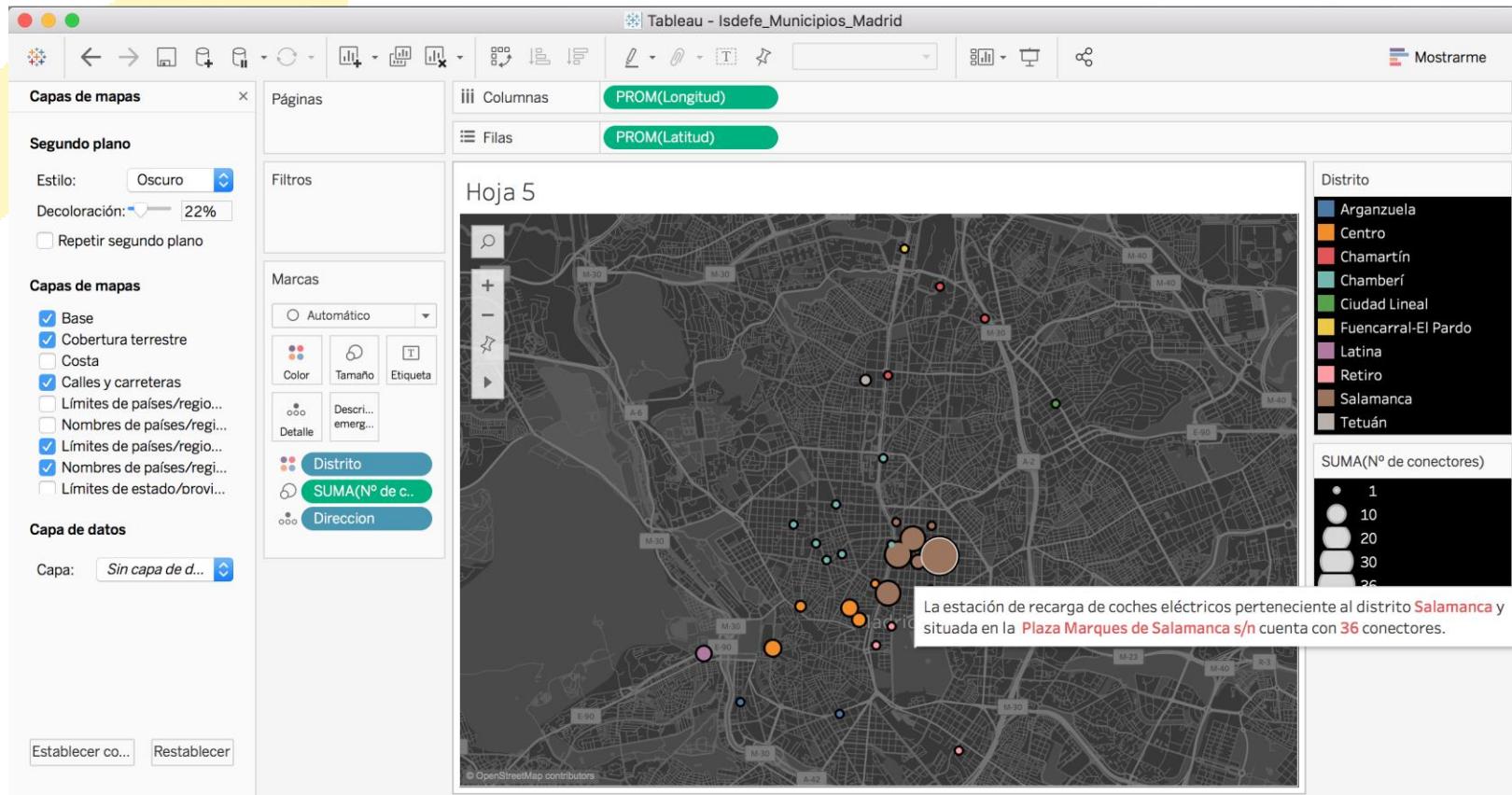
¿Dónde se encuentran las estaciones de recarga? ✓  
¿Cuántos puntos de recarga hay por estación? ✓

## Paso 5: Hacerlo “bonito”



En la Barra de Herramientas disponemos de un menú específico para Mapas. Para hacer el nuestro lo más explicativo posible, añadimos las calles y le damos mayor contraste.

# Paso 6: Autoexplicativo



¡Recordad! Tenemos que adaptar los datos al lenguaje hablado por el receptor. La marca **Descripción Emergente** nos permite poder hacerlo ☺

# Ejercicio

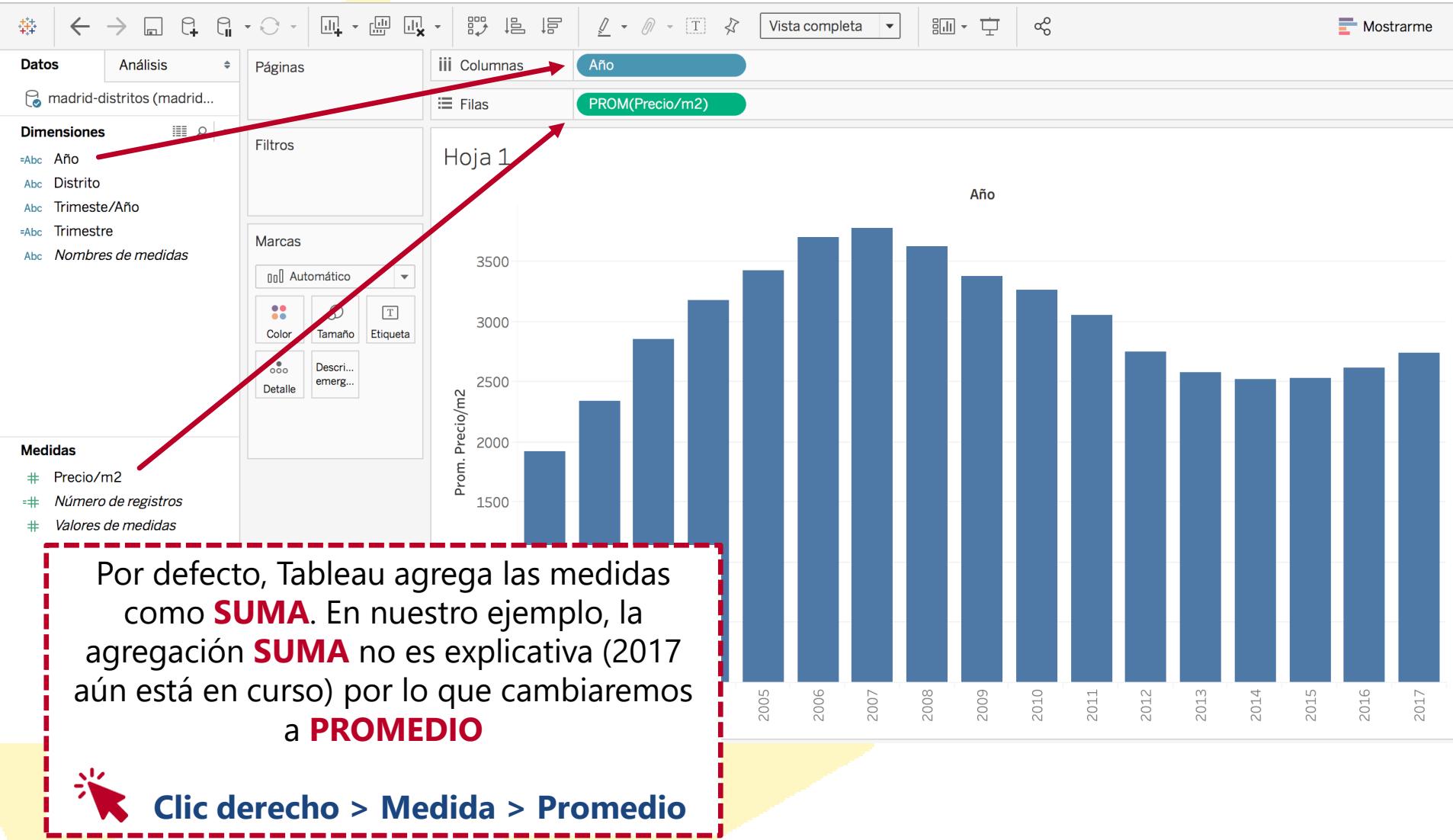


# ¿Qué vamos a ver hoy?

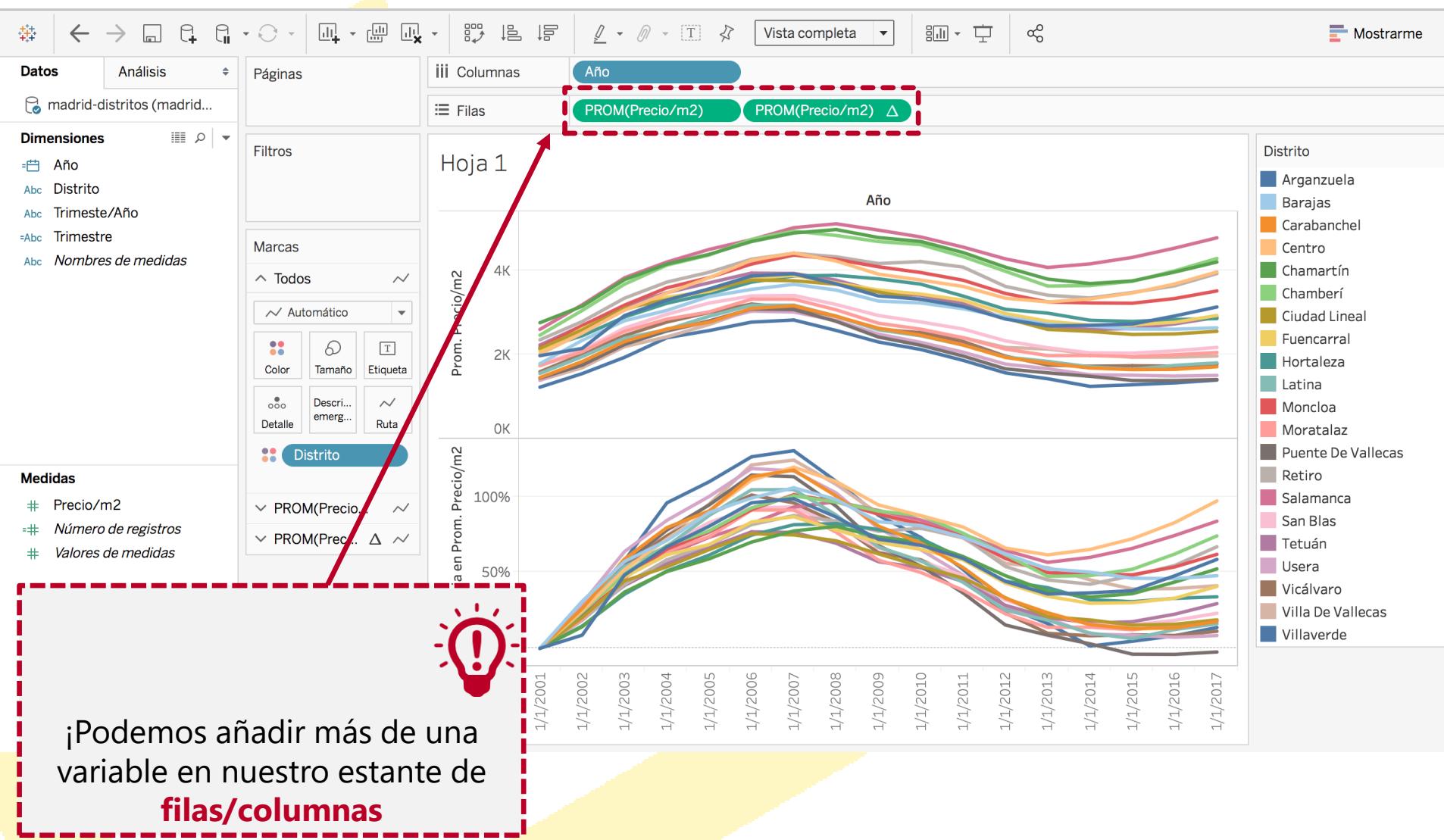
1. Tableau. Interfaz Gráfica
2. Tableau. Arquitectura
3. Tableau para DS
4. Carto

1. Introducción
2. Conexión a los datos
3. Hojas de Trabajo
4. Gráficos predeterminados
5. Cálculos
6. Detección de valores atípicos
7. Combinación de datos
8. Eje Doble
9. Dashboards

# Trabajamos con datos agregados



# ¿Cómo podemos aportar contexto?



# Eficacia visual: Filtros

Tableau interface showing a dashboard with two line charts comparing average price per square meter over time across Madrid districts.

**Dimensions:**

- Año
- Distrito
- Trimestre/Año
- Trimestre
- Nombres de medidas

**Measures:**

- Precio/m<sup>2</sup>

**Filters:**

- Districto (selected)

**Rows:**

- PROM(Precio/m<sup>2</sup>)
- PROM(Precio/m<sup>2</sup>) Δ

**Columns:**

- Año

**Hoja 1 (Top Chart):** Año vs Prom. Precio/m<sup>2</sup>. The Y-axis ranges from 0K to 4K. The chart shows a general upward trend over time, with most districts reaching a peak around 2007-2008 before declining.

**Hoja 2 (Bottom Chart):** Año vs Precio/m<sup>2</sup> Δ. The Y-axis ranges from -100% to 100%. The chart shows the percentage change in price over time, with many districts showing significant peaks and troughs.

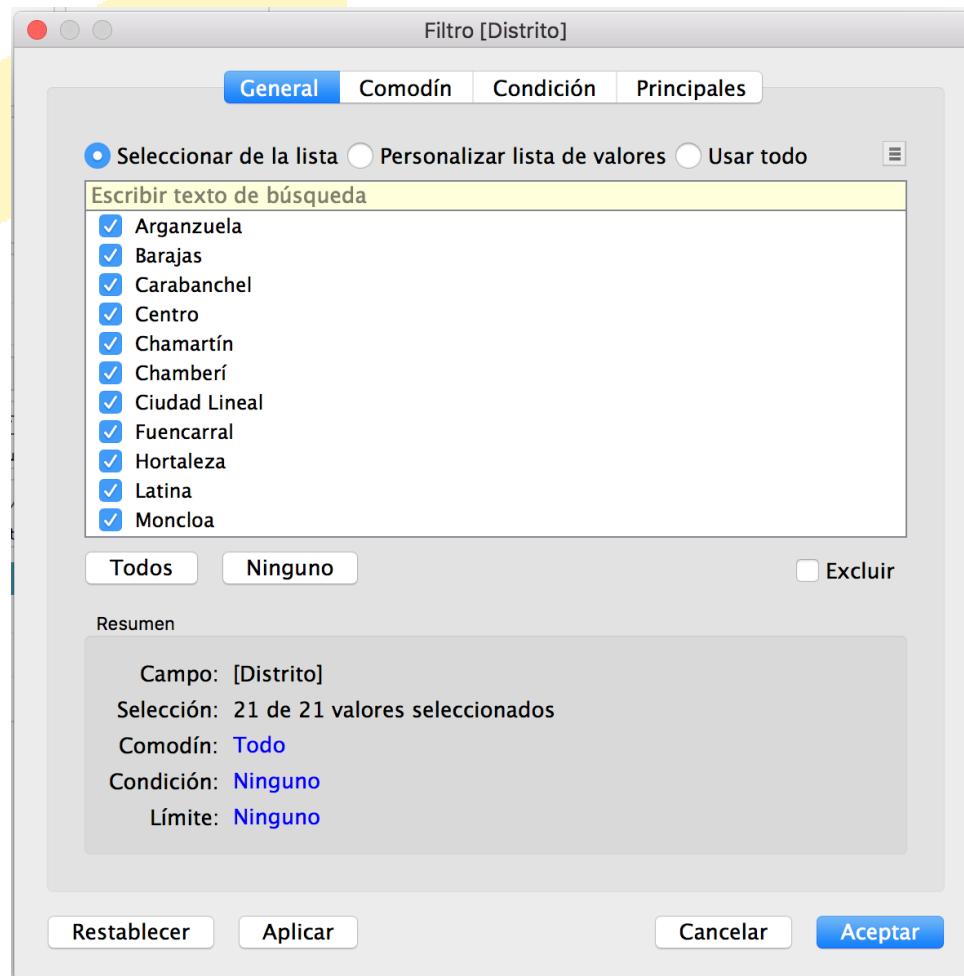
**Legend:**

- Arganzuela
- Barajas
- Carabanchel
- Centro
- Chamartín
- Chamberí
- Ciudad Lineal
- Fuencarral
- Hortaleza
- Latina
- Moncloa
- Moratalaz
- Puente De Vallecas
- Retiro
- Salamanca
- San Blas
- Tetuán
- Usera
- Vicálvaro
- Villa De Vallecas
- Villaverde

**Text Box:**

Tengo un gráfico muy visual y lo he contextualizado, pero la información está enmarañada y apenas se distingue un distrito de otro. ¿Cómo solucionarlo?

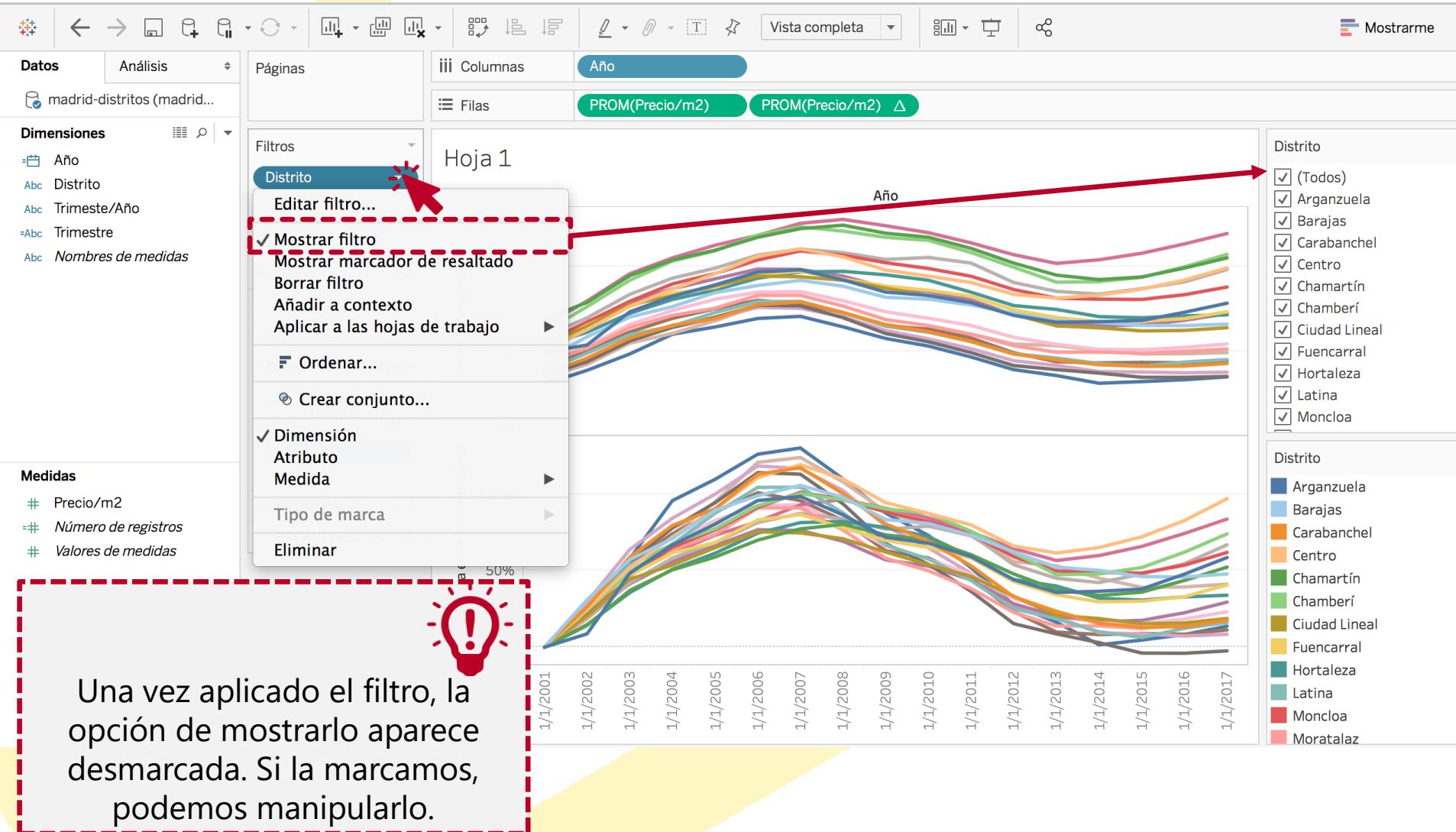
# Eficacia visual: Filtros



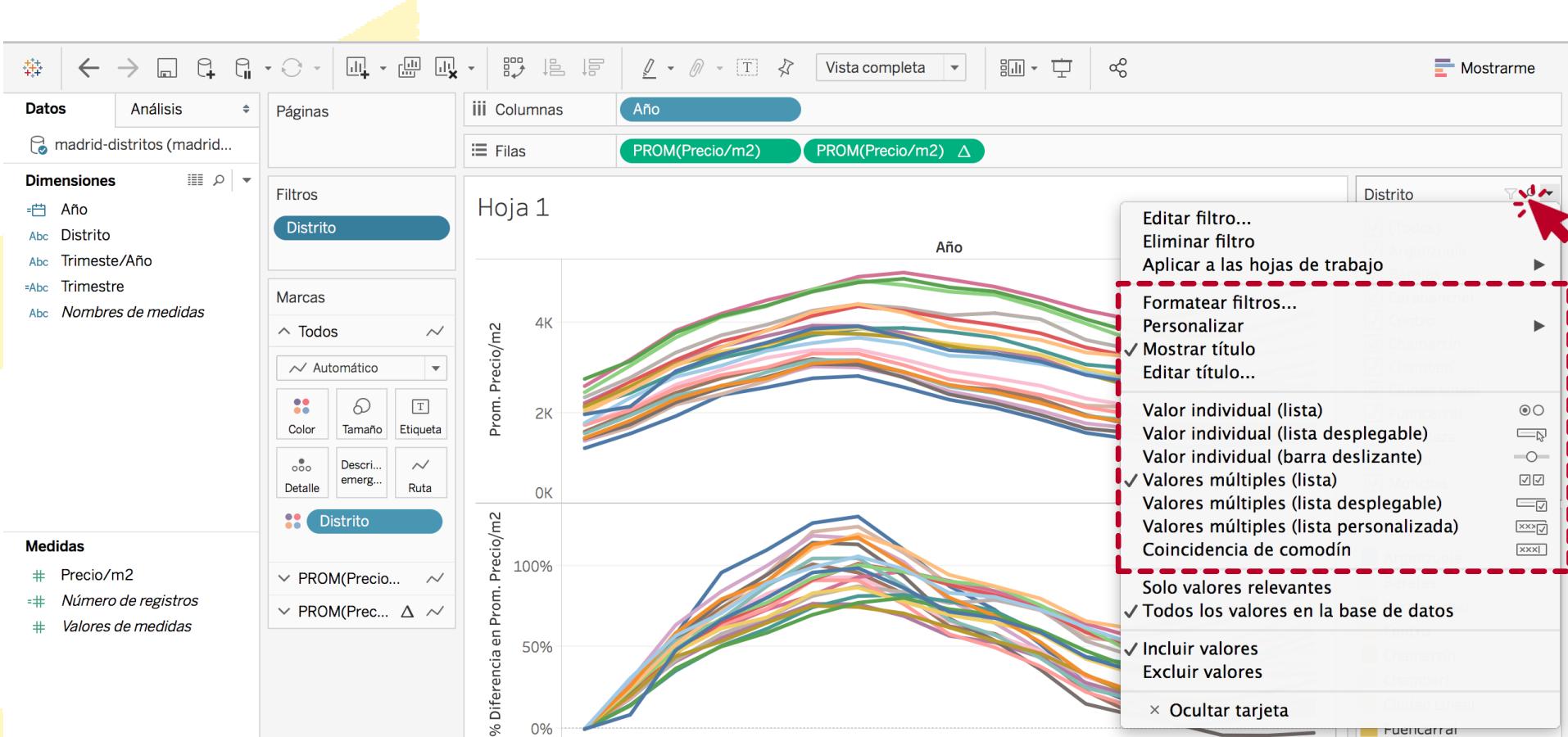
Podemos elegir el tipo de Filtro que queremos aplicar, incluso si de primeras queremos incluir/excluir alguno de los valores que toma la variable.

Por lo general, utilizaremos el filtro general

# Eficacia visual: Filtros

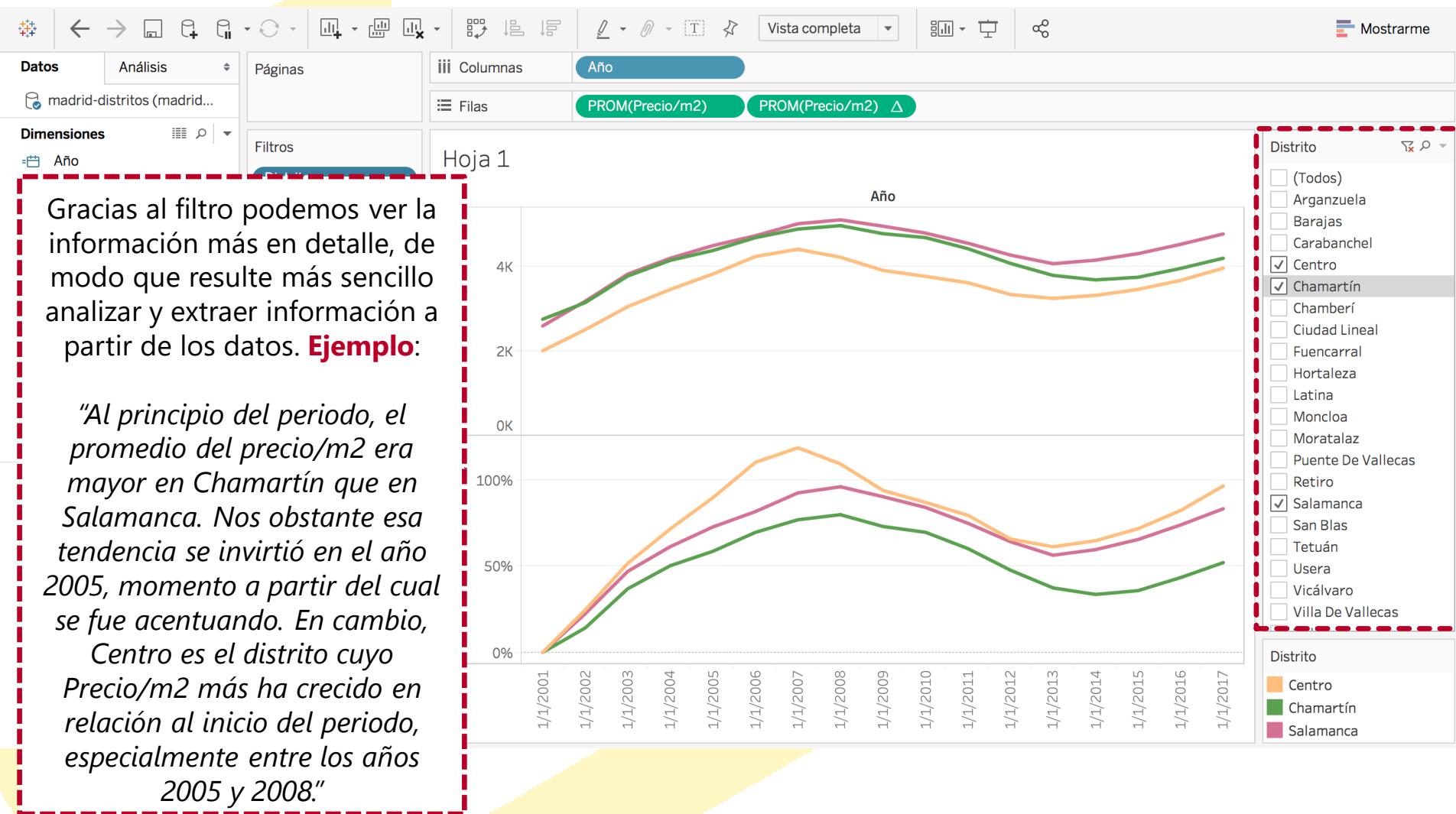


# Eficacia visual: Filtros

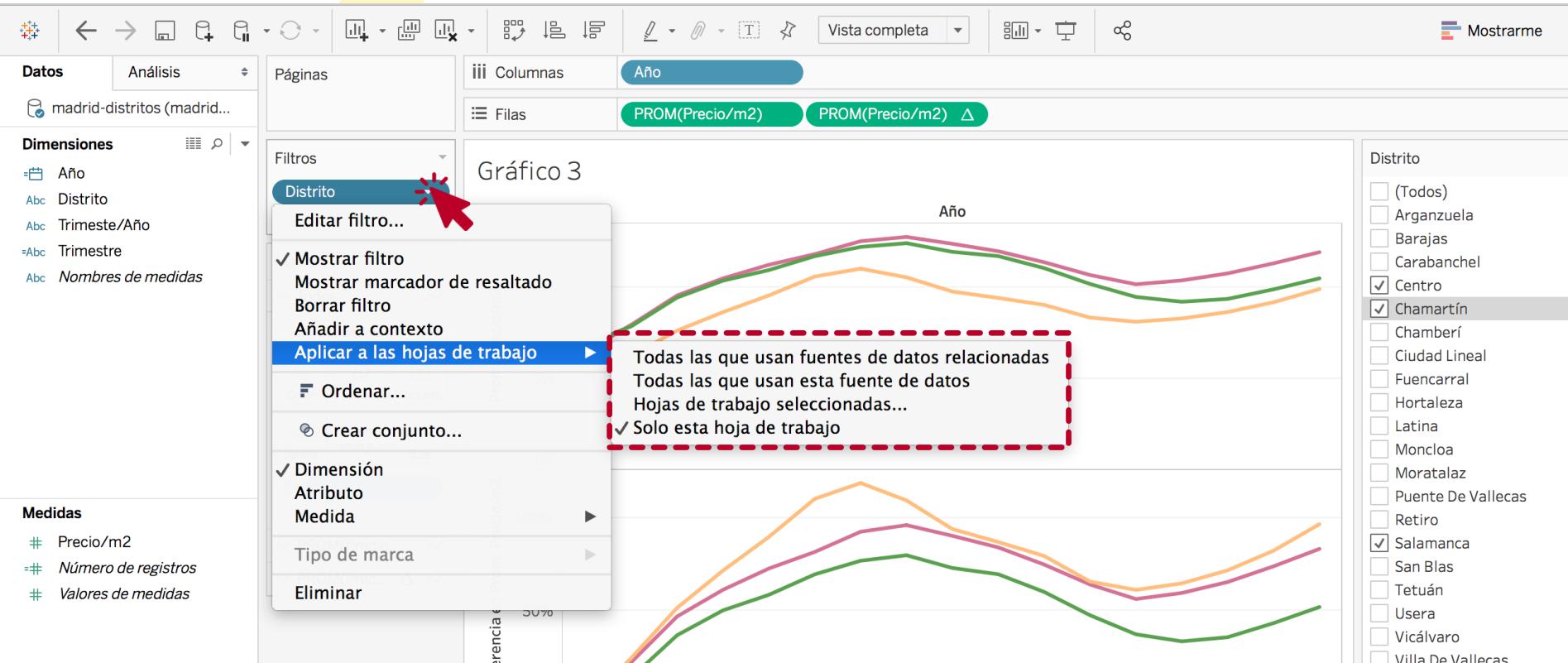


Cuando trabajemos con grandes volúmenes de datos, **previsiblemente necesitaremos formatear el filtro para hacerlo más manipulable**. De ese modo, podremos configurar si queremos que únicamente un valor sea seleccionable, si queremos añadir un buscador al filtro, si preferimos el formato de lista desplegable etc.

# Eficacia visual: Filtros



# Filtros más o menos restrictivos



The screenshot shows the Tableau interface with a context menu open over a filter named "Distrito". The menu items include "Editar filtro...", "Mostrar filtro", "Mostrar marcador de resaltado", "Borrar filtro", "Añadir a contexto", "Aplicar a las hojas de trabajo" (which is selected and highlighted in blue), and "Ordenar...", "Crear conjunto...". Below the menu, a red dashed box highlights the "Aplicar a las hojas de trabajo" option and its sub-options: "Todas las que usan fuentes de datos relacionadas", "Todas las que usan esta fuente de datos", "Hojas de trabajo seleccionadas...", and "Solo esta hoja de trabajo". The "Solo esta hoja de trabajo" option is checked. The main visualization is a line chart titled "Gráfico 3" showing average price per square meter over time for different districts. A legend on the right lists 20 districts, with "Centro" and "Chamartín" checked. The chart shows three distinct peaks corresponding to these districts.

En ocasiones necesitaremos aplicar el filtro únicamente a la hoja en la que estemos trabajando, pero podemos aplicarlo también a toda la fuente de datos e incluso a fuentes de datos relacionadas. Luego lo veremos. 😊

# Ejercicio



# ¿Qué vamos a ver hoy?

1. Tableau. Interfaz Gráfica
2. Tableau. Arquitectura
3. Tableau para DS
4. Carto

1. Introducción
2. Conexión a los datos
3. Hojas de Trabajo
4. Gráficos predeterminados
5. Cálculos
6. Detección de valores atípicos
7. Combinación de datos
8. Eje Doble
9. Dashboards

## Cálculos en Tableau. ¿Qué son?

- ➊ Son el caballo de batalla de Tableau y convierten la herramienta en dinámica y flexible.
- ➋ Permiten llevar a cabo acciones como añadir columnas, "limpiar" los datos, cambiar el tipo de variable etc.
- ➌ En cambio, su dominio requiere un mayor grado de complejidad: algunos cálculos serán muy fáciles mientras que otros nos llevarán más tiempo de aprendizaje y asimilación.



## Cálculos en Tableau. ¿Qué son?

- Los cálculos se implementan dentro de variables conocidas como **Campos Calculados**.
- Para que quede más claro: estos campos calculados se pueden equiparar con la manera en que se utilizan los nombres en Excel; además la sintaxis es parecida. Pueden contener valores fijos o fórmulas que refieran a otras celdas.



Campos  
calculados



Celdas definidas  
por nombres

## Tipos de Cálculos

Número

4+4

Cadena

**"variable -> "VARIABLE"** cambiar el texto o las palabras de una celda

Fecha

**Dec.7, 2013 -> 2013** diseñado para trabajar con fechas fácilmente

Conversión de tipo

**145 -> "145"** cambiar el tipo de variable

Lógico

**IF X > Y THEN "Verdadero"**

Conjunto

**SUM([variable]):** definen formas de juntar (o agregar) a varias filas

Usuario

**"¡Tienes permiso!"** son únicos, porque definen maneras de gestionar cómo la gente se conecta a tus datos

# Cálculos en Tableau. Agregación

 Una de las cosas más “espinosas” de los Cálculos es que pueden ser implementados a **distintos niveles de agregación**, de modo que si el cálculo se define para un determinado nivel de agregación, no funcionará para otros.

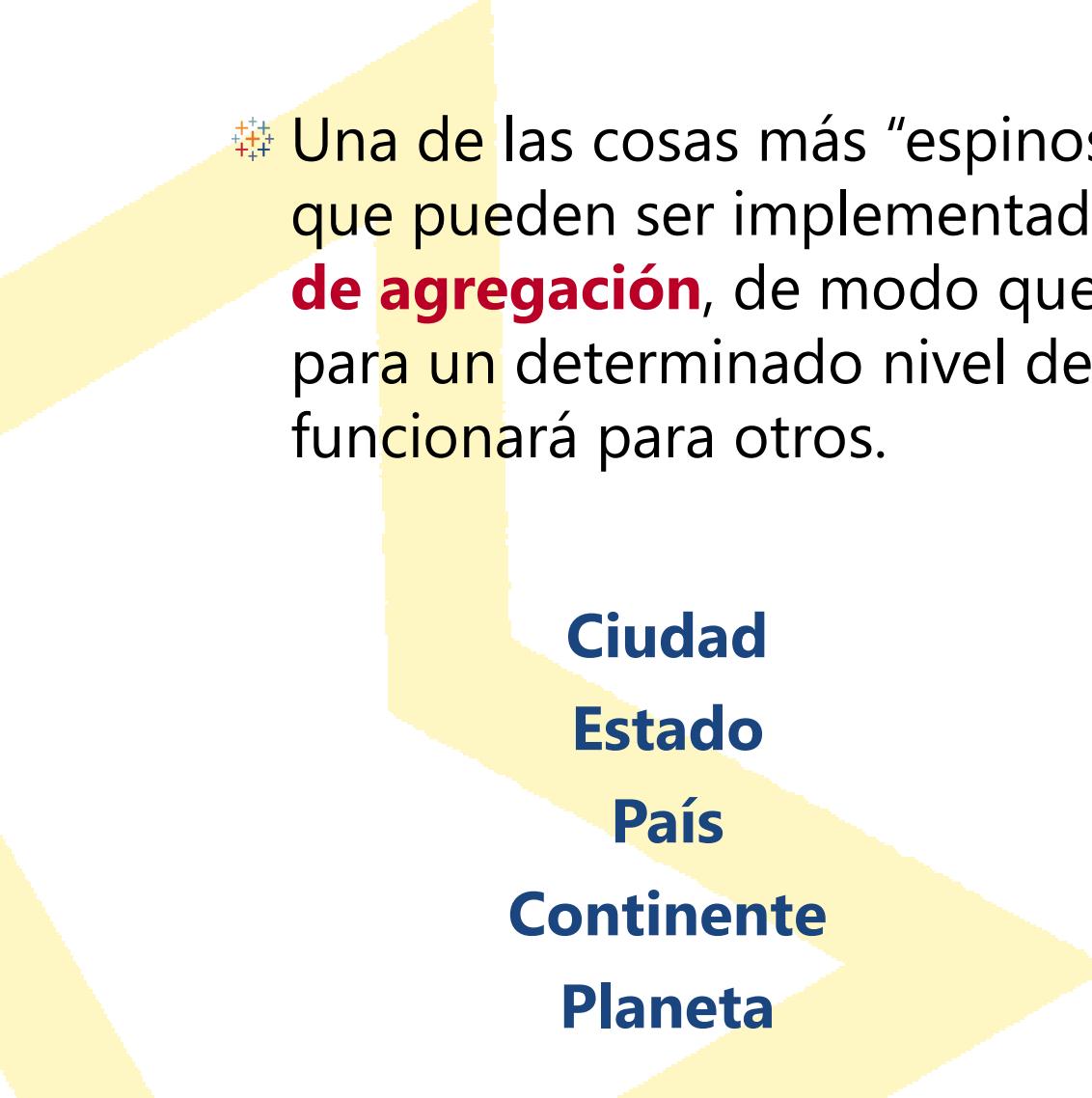


Diagram illustrating levels of aggregation:

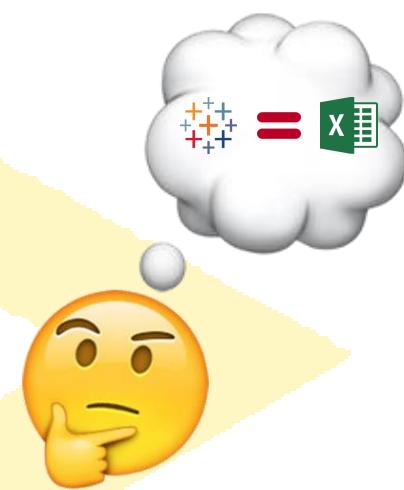
- Ciudad**
- Estado**
- País**
- Continente**
- Planeta**

Portland
Maine
USA
Norteamérica
Tierra

## Cálculos en Tableau. Pistas

La mejor manera de pensar cómo interactúan los cálculos en los distintos niveles de agregación es preguntarse:

**“Si hicéramos este cálculo en Excel, ¿tendría un nuevo número en cada fila, dentro de una nueva columna, o tendría un único número al final de mis columnas iniciales?”**



## Cálculos en Tableau. Pistas

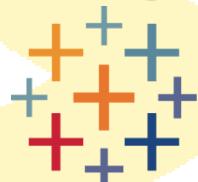
-  **Cálculos de fila ->** se debe usar cuando estemos seguros de que los resultados nos pueden ser útiles en cualquier nivel de detalle
  
-  **Cálculos agregados ->** cálculos que se realizan con múltiples filas a la vez, o cuando las variables ya están agregadas

# Cálculos en Tableau. Pistas



**Cálculos de Tabla ->** Se realizan sólo en los datos que hay dentro de la visualización. Es como si Tableau hiciera cálculos matemáticos en su cabeza. Son extremadamente potentes, pero pueden ser de las partes más difíciles de aprender a dominar en Tableau.

$$x + y = z$$



# ¿Cómo se escriben los Cálculos?

**Título:** Siempre hay que tener uno, y debe ser lo más descriptivo posible.



The screenshot shows the Tableau calculation editor. On the left, there is a text input field labeled "Cálculo1" with a red dashed border. To the right of the input field is a "X" button. Below the input field is a button labeled "Aplicar" (Apply) and a green "Aceptar" (Accept) button. To the right of the main area is a vertical list of available calculations, each preceded by a blue circular icon with a white sunburst pattern. The list includes: Todos, ABS, ACOS, AND, ASCII, ASIN, ATAN, ATAN2, ATTR, AVG, CASE, and CEILING. A blue arrow points from the text "Área donde se escribe el cálculo" to the input field. Another blue arrow points from the text "Cálculos disponibles" to the list of functions.

**Área donde se escribe el cálculo**

Cálculo1

X

Aplicar    Aceptar

Todos

Escribir texto de búsqueda

ABS (número)

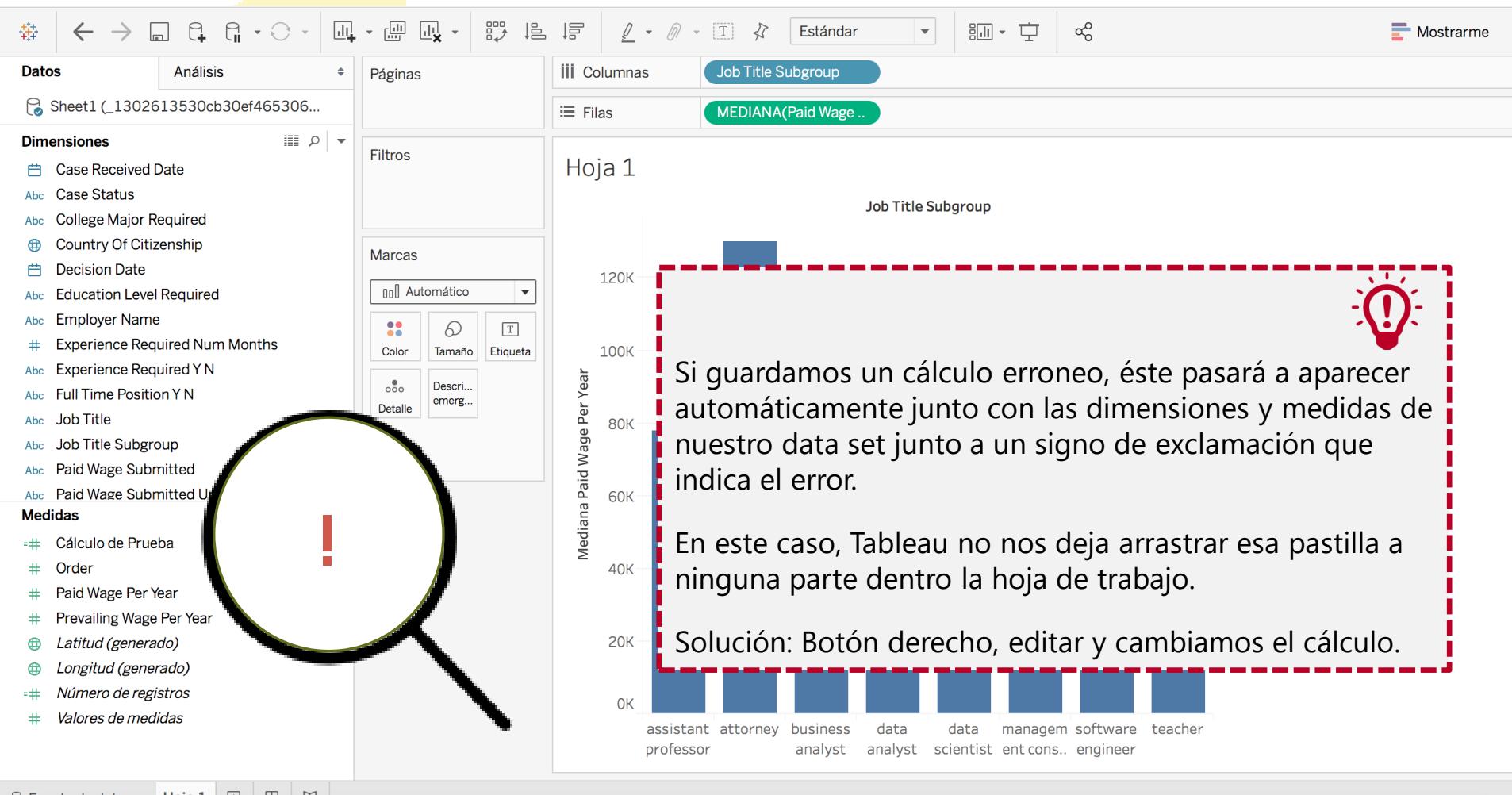
Indica el valor absoluto del número dado.

Ejemplo:  $\text{ABS}(-7) = 7$

ABS  
ACOS  
AND  
ASCII  
ASIN  
ATAN  
ATAN2  
ATTR  
AVG  
CASE  
CEILING

**Cálculos disponibles:** Una de las cosas buenas de Tableau es que te dice lo que se necesita escribir dentro de cada una de las funciones. Abriendo el desplegable, podemos ver las distintas categorías existentes para estos cálculos.

# Errores en los Cálculos



The screenshot shows the Tableau interface with a calculated field error. In the top right corner, there is a red dashed box containing a lightbulb icon with an exclamation mark, indicating an error. The error message inside the box reads: "Si guardamos un cálculo erróneo, éste pasará a aparecer automáticamente junto con las dimensiones y medidas de nuestro data set junto a un signo de exclamación que indica el error." Below this, another message states: "En este caso, Tableau no nos deja arrastrar esa pastilla a ninguna parte dentro la hoja de trabajo." At the bottom, a solution is provided: "Solución: Botón derecho, editar y cambiamos el cálculo." To the left of the error message, there is a magnifying glass icon with an exclamation mark inside it.

Datos Análisis

Páginas Job Title Subgroup

Sheet1 (\_1302613530cb30ef465306...)

Dimensions

- Case Received Date
- Case Status
- College Major Required
- Country Of Citizenship
- Decision Date
- Education Level Required
- Employer Name
- Experience Required Num Months
- Experience Required Y N
- Full Time Position Y N
- Job Title
- Job Title Subgroup
- Paid Wage Submitted
- Paid Wage Submitted U

Medidas

- Cálculo de Prueba
- Order
- Paid Wage Per Year
- Prevailing Wage Per Year
- Latitud (generado)
- Longitud (generado)
- Número de registros
- Valores de medidas

Columnas Job Title Subgroup

Filas MEDIANA(Paid Wage ...)

Hoja 1

Job Title Subgroup

Mediana Paid Wage Per Year

assistant professor attorney business analyst data analyst data scientist managem ent cons... software engineer teacher

! Si guardamos un cálculo erróneo, éste pasará a aparecer automáticamente junto con las dimensiones y medidas de nuestro data set junto a un signo de exclamación que indica el error.

En este caso, Tableau no nos deja arrastrar esa pastilla a ninguna parte dentro la hoja de trabajo.

Solución: Botón derecho, editar y cambiamos el cálculo.

# Gráficos: Dispersion

**Para seguir las explicaciones, os invito a que:**

- 1) Creéis una nueva hoja de trabajo y la llaméis “En busca de outliers”.
- 2) Llevéis la medida **Paid Wage per Year** al estante filas (agregado como Mediana) y la dimensión **Job Title Subgroup** al estante columnas.
- 3) Ordenéis los resultados (gráfico de barras) de mayor a menor.

**Pregunta: ¿por qué nos interesa encontrar “Outliers”?**

# ¿Qué vamos a ver hoy?

1. Tableau. Interfaz Gráfica
2. Tableau. Arquitectura
3. Tableau para DS
4. Carto

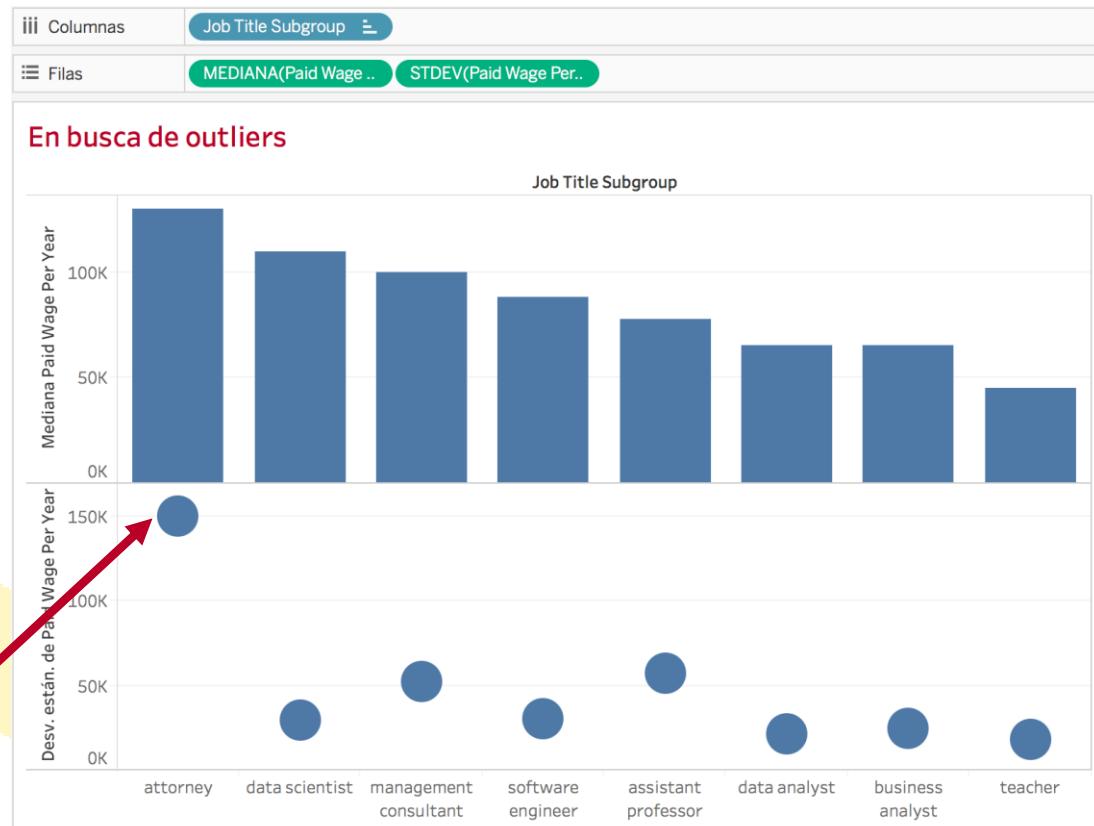
1. Introducción
2. Conexión a los datos
3. Hojas de Trabajo
4. Gráficos predeterminados
5. Cálculos
6. Detección de valores atípicos
7. Combinación de datos
8. Eje Doble
9. Dashboards

## Gráficos: Dispersión

Para intuir de un vistazo rápido si existen valores atípicos en nuestros datos:

1 ) Arrastra la medida **Paid Wage per Year** (agregada como Desviación Típica) al estante filas.

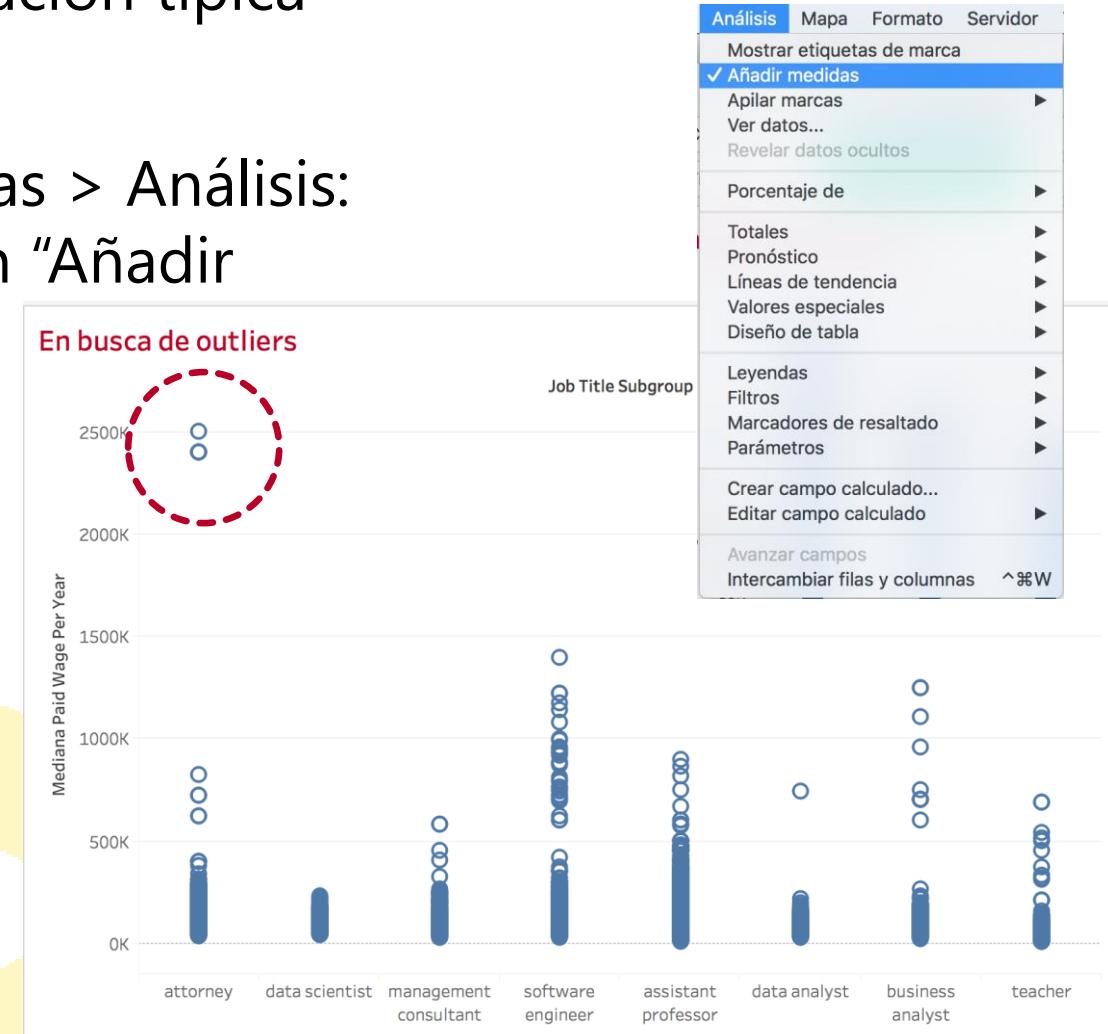
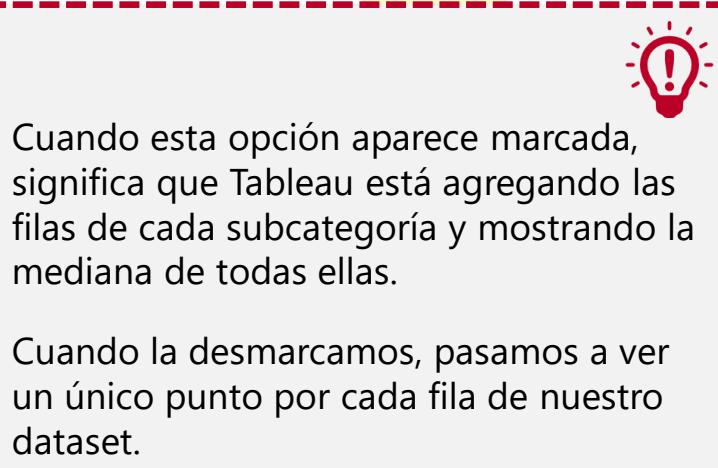
2) Cambia la forma de este último gráfico introducido a círculo. **¿Qué veis?**



## Gráficos: Dispersión

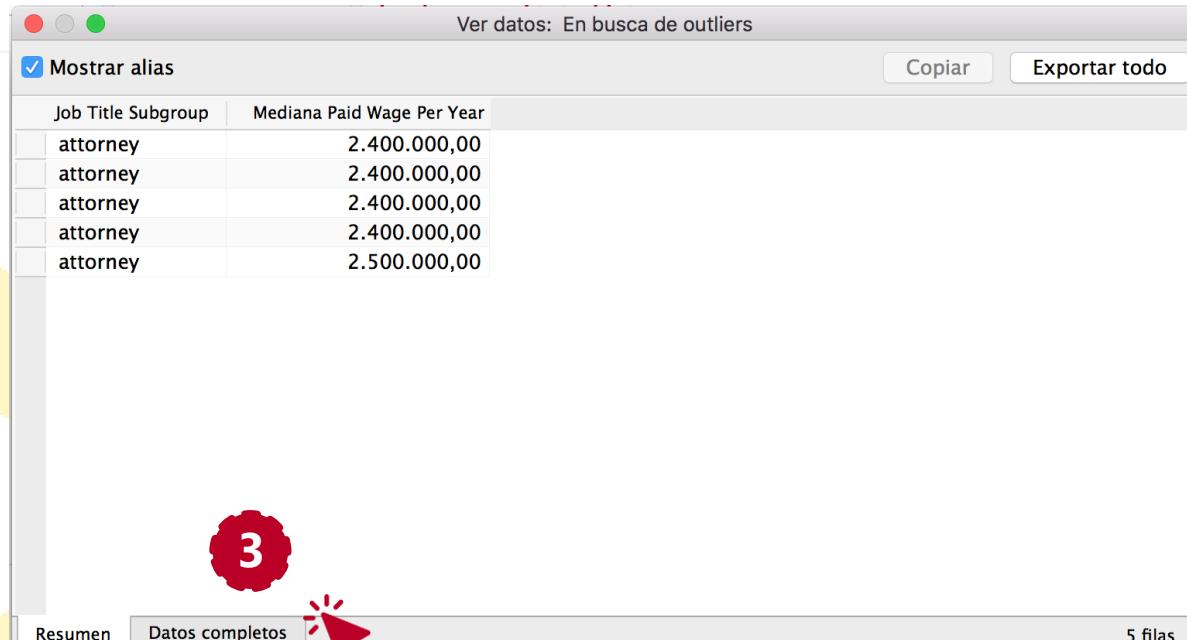
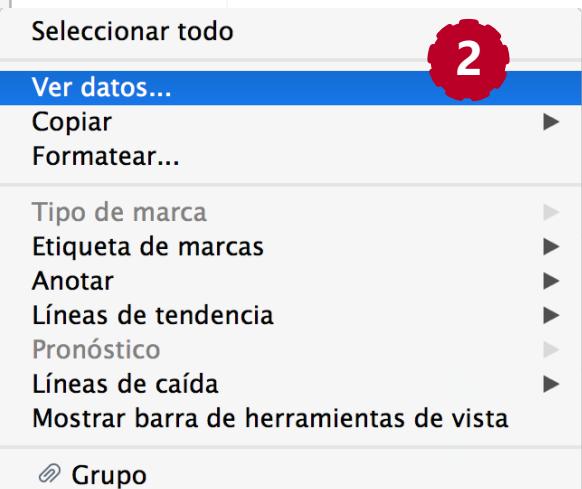
1 ) Eliminamos la desviación típica del gráfico.

2) Barra de Herramientas > Análisis: desmarcamos la opción “Añadir Medidas”.



## Indagar en los datos

### En busca de outliers



The screenshot shows a 'Ver datos' window titled 'Ver datos: En busca de outliers'. It has a checkbox for 'Mostrar alias' which is checked. Below is a table with two columns: 'Job Title Subgroup' and 'Mediana Paid Wage Per Year'. The data consists of five rows, all of which are 'attorney' with wages ranging from 2.400.000,00 to 2.500.000,00. The bottom of the window shows tabs for 'Resumen' and 'Datos completos', with 'Datos completos' being the active tab. A red circle with the number '3' is placed over the 'Datos completos' tab, and a red cursor arrow is pointing towards it.

Job Title Subgroup	Mediana Paid Wage Per Year
attorney	2.400.000,00
attorney	2.500.000,00

- 1) Seleccionamos el área donde se encuentran nuestros valores atípicos
- 2) Botón Derecho > Ver Datos
- 3) Clic en Datos Completos

# ¿Qué nos llama la atención?

Ver datos: En busca de outliers

5 filas  Mostrar alias  Mostrar todos los campos Copiar Exportar todo

Case Number	Case Received Date	Case Status	College Major Required	Country Of Citizenship	Decision Date	Education Level Required	Employer Name	Estados Interesantes	Expo
I-200-13081-043757	28/3/2013	certified	NA	NA	3/4/2013	NA	HUGHES HUBBARD & REED LLP	Viviría aquí	
I-203-14021-228428	5/2/2014	withdrawn	NA	NA	5/2/2014	NA	WINSTON & STRAWN LLP	No viviría aquí	
I-200-13077-839340	18/3/2013	denied	NA	NA	22/3/2013	NA	HUGHES HUBBARD & REED LLP	No viviría aquí	
I-200-13081-136705	22/3/2013	denied	NA	NA	28/3/2013	NA	HUGHES HUBBARD & REED LLP	No viviría aquí	
I-200-13087-022976	28/3/2013	certified	NA	NA	3/4/2013	NA	HUGHES HUBBARD & REED LLP	No viviría aquí	

Resumen Datos completos 5 filas



Dos de ellas se presentaron el mismo día.



Cuatro de las cinco ofertas provienen de la misma empresa. ¿De verdad esa empresa está dispuesta a pagar más de dos millones de \$ a sus empleados? ¿O alguien ha cometido un error a la hora de presentar la oferta?

## En base a los datos, ¿a qué estado de EEUU te irías a vivir?



**No tenemos info suficiente para responder a esa pregunta**

Por ejemplo, un salario de 75.000\$ no es lo mismo en New Hampshire que en Texas o San Francisco.

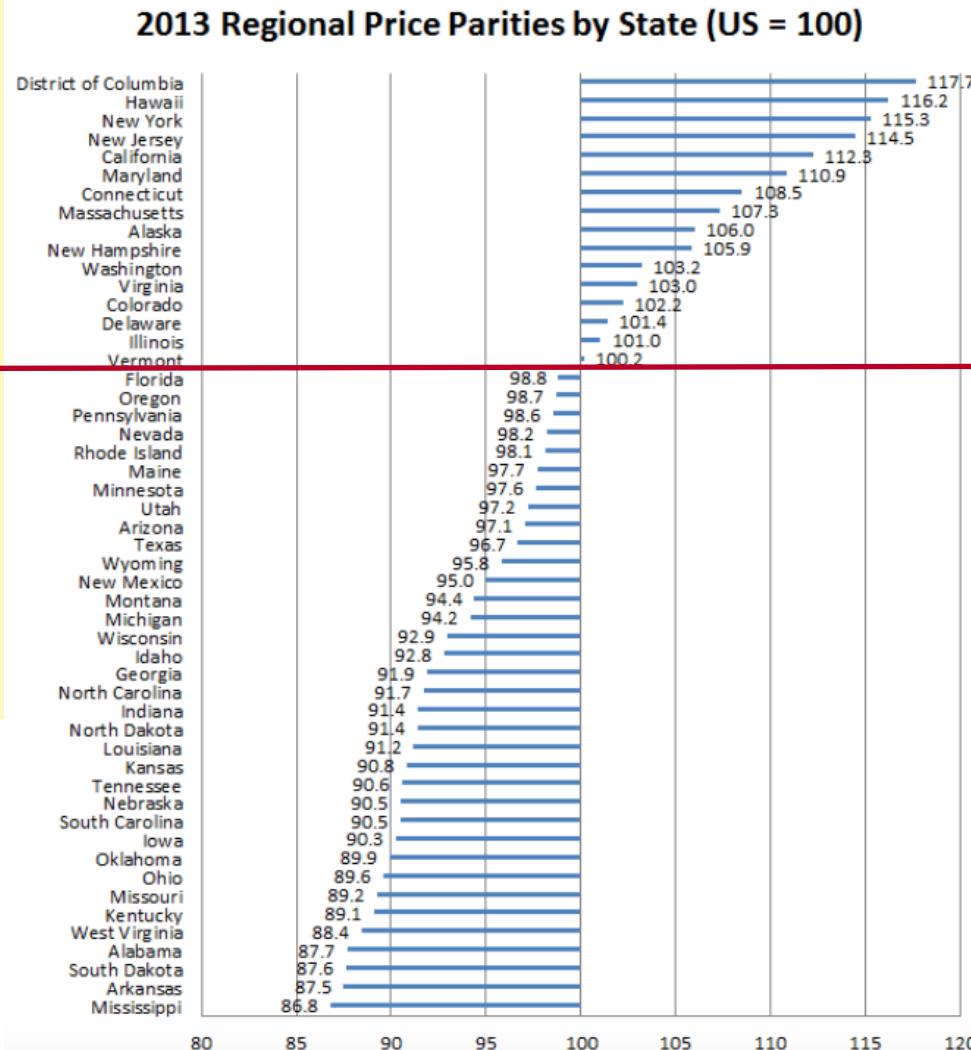
Necesitaríamos ajustar los salarios al coste de la vida.



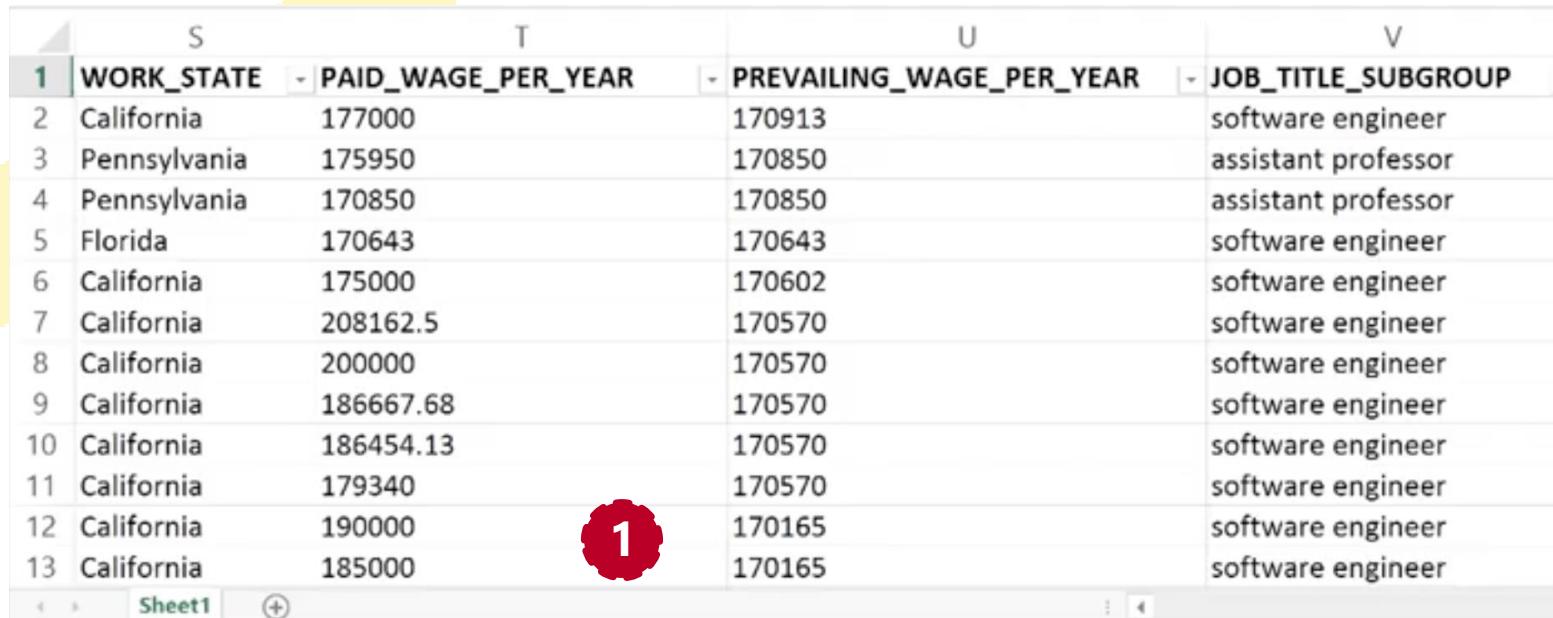
## ¿Qué vamos a ver?

- 
- The slide features a large, stylized yellow chevron graphic on the left side, pointing towards the right where the content is listed. The text is organized into two columns separated by a vertical bracket.
- |                              |   |
|------------------------------|---|
| 1. Tableau. Interfaz Gráfica | 1. Introducción<br>2. Conexión a los datos<br>3. Hojas de Trabajo<br>4. Gráficos predeterminados<br>5. Cálculos<br>6. Detección de valores atípicos<br>7. Combinación de datos<br>8. Eje Doble<br>9. Dashboards |
| 2. Tableau. Arquitectura     |   |
| 3. Tableau para DS           |   |
| 4. Carto                     |   |

# Carácterísticas del nuevo dataset



# Problemas entre datasets



1	S	T	U	V
1	WORK_STATE	PAID_WAGE_PER_YEAR	PREVAILING_WAGE_PER_YEAR	JOB_TITLE_SUBGROUP
2	California	177000	170913	software engineer
3	Pennsylvania	175950	170850	assistant professor
4	Pennsylvania	170850	170850	assistant professor
5	Florida	170643	170643	software engineer
6	California	175000	170602	software engineer
7	California	208162.5	170570	software engineer
8	California	200000	170570	software engineer
9	California	186667.68	170570	software engineer
10	California	186454.13	170570	software engineer
11	California	179340	170570	software engineer
12	California	190000	170165	software engineer
13	California	185000	170165	software engineer

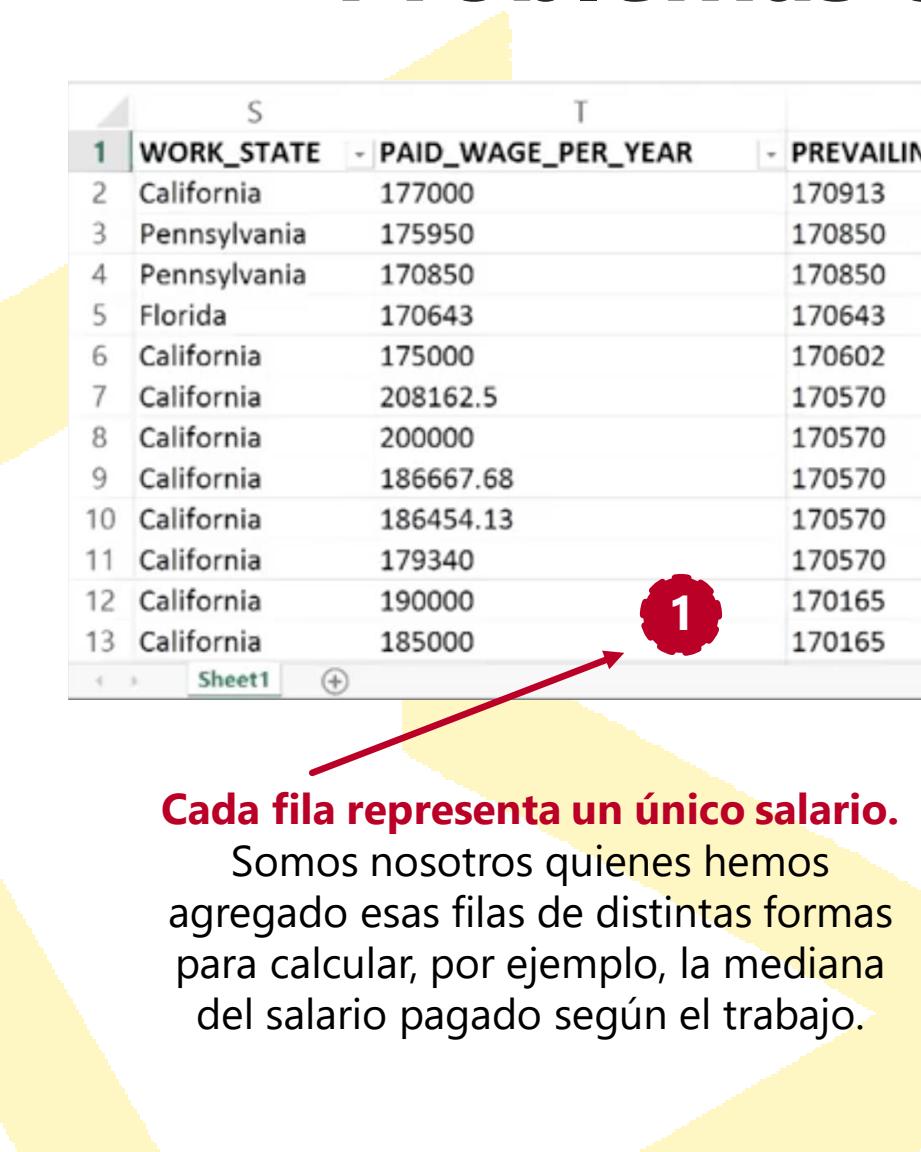
Nuestros datasets tienen **distintos niveles de granularidad**



1	A	B
1	State	Price Parity
2	Alabama	87.7
3	Alaska	106.0
4	Arizona	97.1
5	Arkansas	87.5
6	California	112.3
7	Colorado	102.2
8	Connecticut	108.5

2

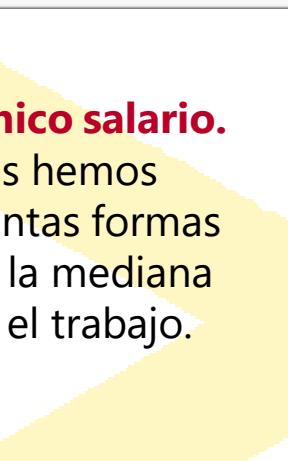
# Problemas entre datasets



1	WORK_STATE	PAID_WAGE_PER_YEAR	PREVAILING_WAGE_PER_YEAR	JOB_TITLE_SUBGROUP
2	California	177000	170913	software engineer
3	Pennsylvania	175950	170850	assistant professor
4	Pennsylvania	170850	170850	assistant professor
5	Florida	170643	170643	software engineer
6	California	175000	170602	software engineer
7	California	208162.5	170570	software engineer
8	California	200000	170570	software engineer
9	California	186667.68	170570	software engineer
10	California	186454.13	170570	software engineer
11	California	179340	170570	software engineer
12	California	190000	170165	software engineer
13	California	185000	170165	software engineer

**Cada fila representa un único salario.**

Somos nosotros quienes hemos agregado esas filas de distintas formas para calcular, por ejemplo, la mediana del salario pagado según el trabajo.



A	B	C
State	Price Parity	
2 Alabama	87.7	
3 Alaska	106.0	
4 Arizona	97.1	
5 Arkansas	87.5	
6 California	112.3	
7 Colorado	102.2	
8 Connecticut	108.5	

# Problemas entre datasets

1	WORK_STATE	PAID_WAGE_PER_YEAR	PREVAILING_WAGE_PER_YEAR	JOB_TITLE_SUBGROUP
2	California	177000	170913	software engineer
3	Pennsylvania	175950	170850	assistant professor
4	Pennsylvania	170850	170850	assistant professor
5	Florida	170643	170643	software engineer
6	California	175000	170602	software engineer
7	California	208162.5	170570	software engineer
8	California	200000	170570	software engineer
9	California	186667.68	170570	software engineer
10	California	186454.13	170570	software engineer
11	California	179340	170570	software engineer
12	California	190000	170165	software engineer
13	California	185000	170165	software engineer

Cada fila representa un único Estado Americano.

Tableau nos va ayudar a solucionar el problema a través de la combinación de datos o Data Blending.

A	B	C
State	Price Parity	
Alabama	87.7	
Alaska	106.0	
Arizona	97.1	
Arkansas	87.5	
California	112.3	
Colorado	102.2	
Connecticut	108.5	

## Data blending

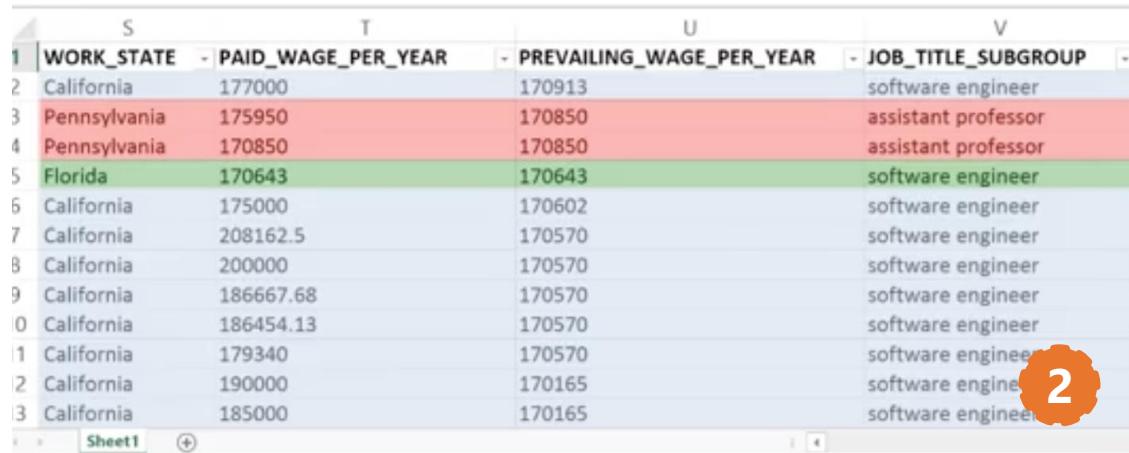
*"Es la manera de combinar dos datasets cuando éstos tienen distintos niveles de agregación. Lo único que necesita Tableau es un campo que sea común a ambos datasets."*

En nuestro ejemplo, ese campo común va a ser la variable [Work State](#), que hace referencia al nombre del Estado.

# Data Blending. Eficiencia

Desde el punto de vista computacional, la forma más eficiente de realizar la combinación **es agregar primero el dataset más grande** (con mayor granularidad) al nivel de agregación del campo común, definido en el dataset menos granular.

Esto es así porque Tableau tiene que crear una tabla temporal y almacenarla en su memoria.



S	T	U	V
WORK_STATE	PAID_WAGE_PER_YEAR	PREVAILING_WAGE_PER_YEAR	JOB_TITLE_SUBGROUP
California	177000	170913	software engineer
Pennsylvania	175950	170850	assistant professor
Pennsylvania	170850	170850	assistant professor
Florida	170643	170643	software engineer
California	175000	170602	software engineer
California	208162.5	170570	software engineer
California	200000	170570	software engineer
California	186667.68	170570	software engineer
California	186454.13	170570	software engineer
California	179340	170570	software engineer
California	190000	170165	software engineer
California	185000	170165	software engineer



A	B	C
State	Price Parity	
Alabama	87.7	1
Alaska	106.0	
Arizona	97.1	
Arkansas	87.5	
California	112.3	
Colorado	102.2	
Connecticut	108.5	
Delaware	101.4	

# Data Blending. Eficiencia

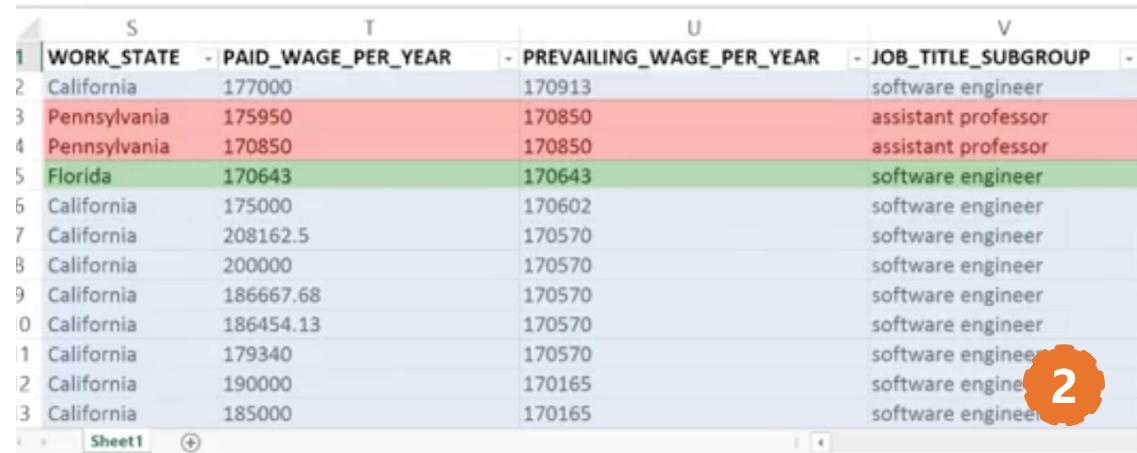
Esta tabla temporal será más pequeña si agregamos el dataset grande primero y sólo entonces combinamos las dos tablas.

## ¿Si fuera a la inversa?

Habría que llenar cada fila del dataset grande con el mismo valor repetido existente en la tabla pequeña.

**Dataset primario:** el que tenga menor grado de granularidad

**Dataset secundario:** el que tenga mayor grado de granularidad



S	T	U	V
WORK_STATE	PAID_WAGE_PER_YEAR	PREVAILING_WAGE_PER_YEAR	JOB_TITLE_SUBGROUP
California	177000	170913	software engineer
Pennsylvania	175950	170850	assistant professor
Pennsylvania	170850	170850	assistant professor
Florida	170643	170643	software engineer
California	175000	170602	software engineer
California	208162.5	170570	software engineer
California	200000	170570	software engineer
California	186667.68	170570	software engineer
California	186454.13	170570	software engineer
California	179340	170570	software engineer
California	190000	170165	software engineer
California	185000	170165	software engineer

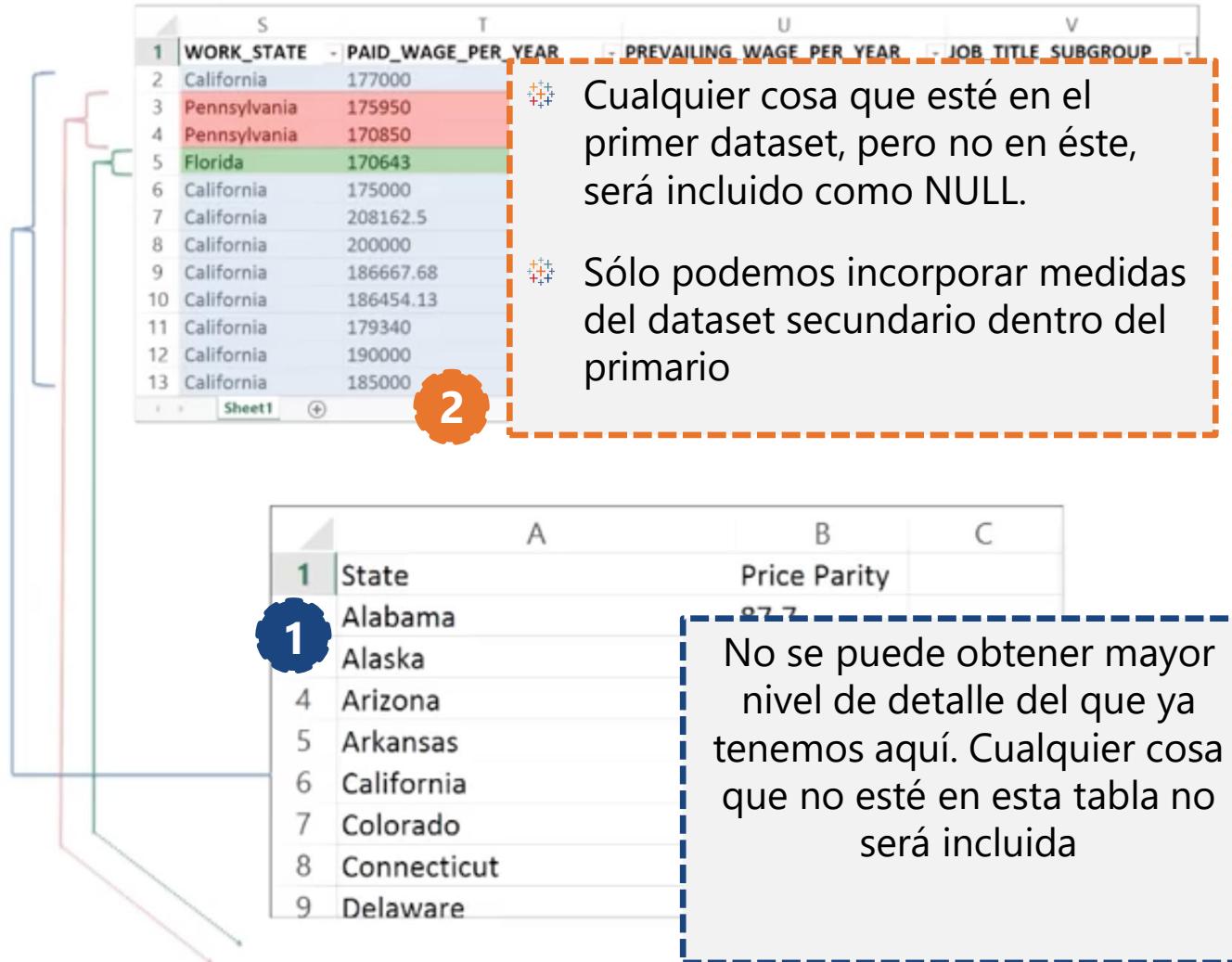


A	B	C
State	Price Parity	
Alabama	87.7	1
Alaska	106.0	
Arizona	97.1	
Arkansas	87.5	
California	112.3	
Colorado	102.2	
Connecticut	108.5	
Delaware	101.4	

# Data Blending. Consecuencias



La combinación se llevará cada vez que pintemos un gráfico



1 WORK\_STATE - PAID\_WAGE\_PER\_YEAR - PREVAILING\_WAGE\_PER\_YEAR - JOB\_TITLE\_SUBGROUP

2 California 177000  
3 Pennsylvania 175950  
4 Pennsylvania 170850  
5 Florida 170643  
6 California 175000  
7 California 208162.5  
8 California 200000  
9 California 186667.68  
10 California 186454.13  
11 California 179340  
12 California 190000  
13 California 185000

Sheet1

1 State

2 Alabama  
3 Alaska  
4 Arizona  
5 Arkansas  
6 California  
7 Colorado  
8 Connecticut  
9 Delaware

B Price Parity

C

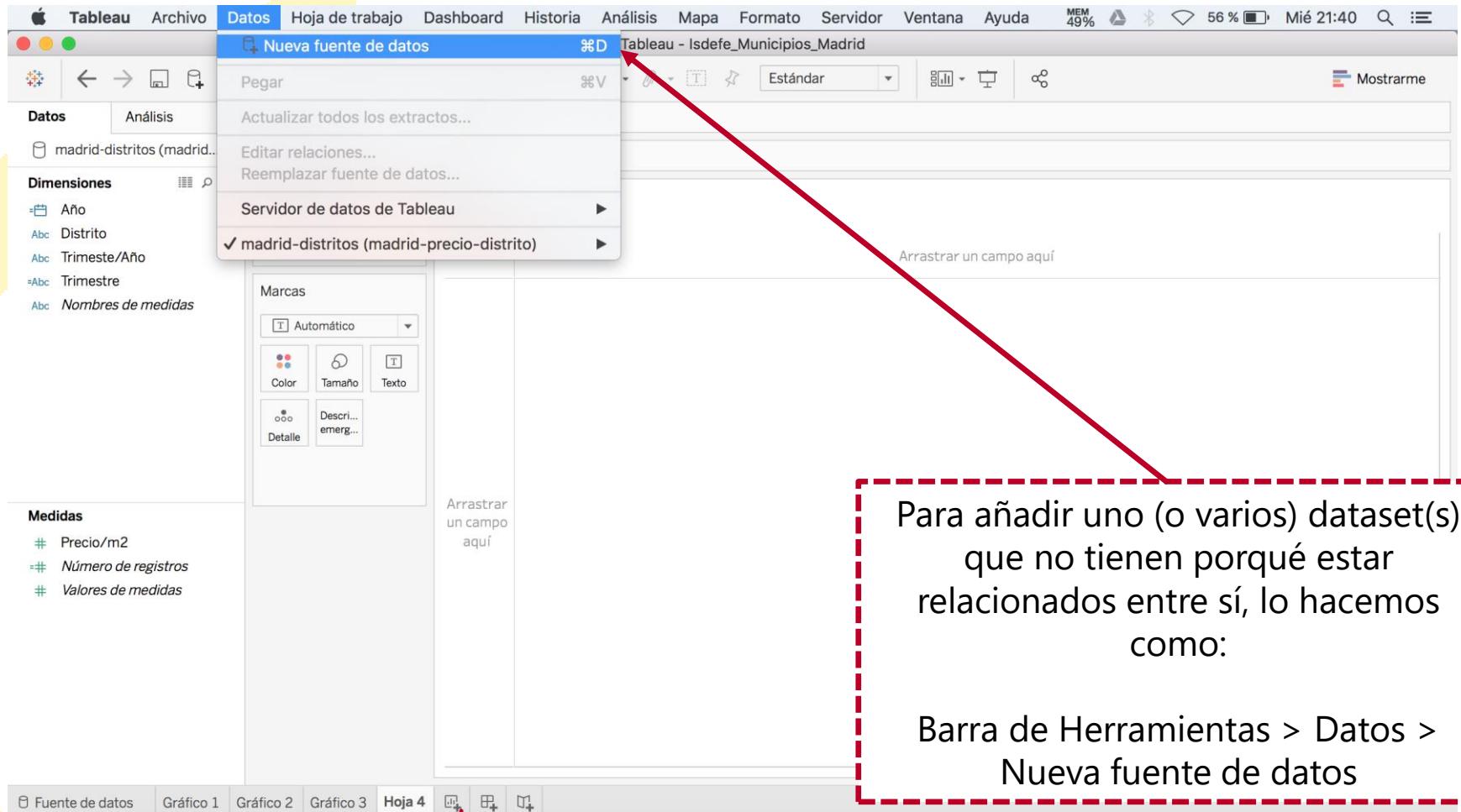
2

Cualquier cosa que esté en el primer dataset, pero no en éste, será incluido como NULL.

Sólo podemos incorporar medidas del dataset secundario dentro del primario

No se puede obtener mayor nivel de detalle del que ya tenemos aquí. Cualquier cosa que no esté en esta tabla no será incluida

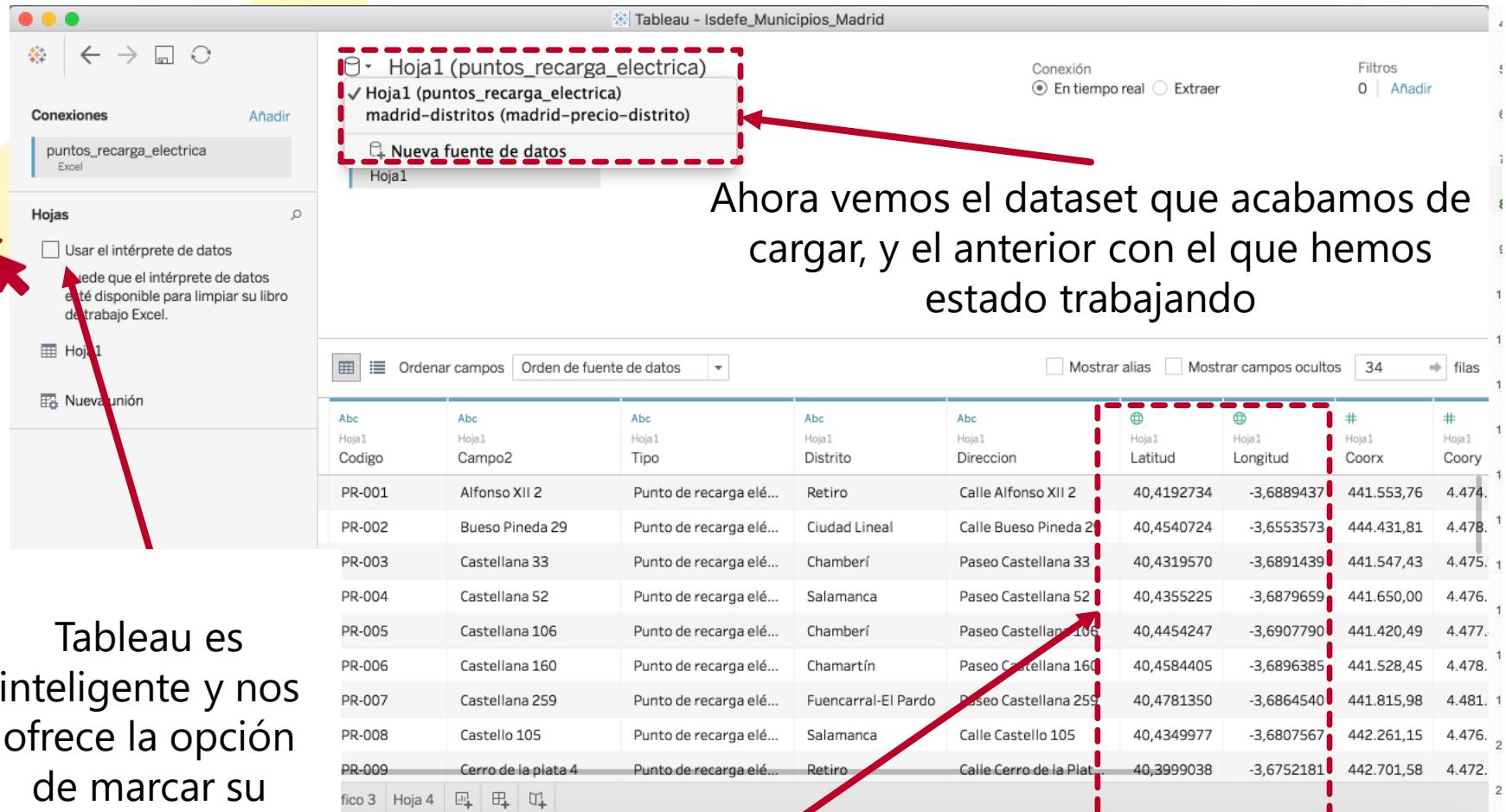
## ¿Cómo añadimos el segundo dataset?



Para añadir uno (o varios) dataset(s) que no tienen porque estar relacionados entre sí, lo hacemos como:

Barra de Herramientas > Datos >  
Nueva fuente de datos

# ¿Cómo añadimos el segundo dataset?



Conexiones Añadir  
puntos\_recarga\_electrica Excel  
Hojas  
Usar el intérprete de datos  
Necesita que el intérprete de datos esté disponible para limpiar su libro de trabajo Excel.  
Hoja1  
Nueva unión

Hoja1 (puntos\_recarga\_electrica)  
✓ Hoja1 (puntos\_recarga\_electrica)  
madrid-distritos (madrid-precio-districto)  
Nueva fuente de datos  
Hoja1

Conexión En tiempo real Extraer  
Filtros 0 Añadir

Abc Hoja1 Codigo	Abc Hoja1 Campo2	Abc Hoja1 Tipo	Abc Hoja1 Distrito	Abc Hoja1 Direccion	Hoja1 Latitud	Hoja1 Longitud	# Hoja1 Coorx	# Hoja1 Coory
PR-001	Alfonso XII 2	Punto de recarga elé...	Retiro	Calle Alfonso XII 2	40,4192734	-3,6889437	441.553,76	4.474.
PR-002	Bueso Pineda 29	Punto de recarga elé...	Ciudad Lineal	Calle Bueso Pineda 29	40,4540724	-3,6553573	444.431,81	4.478.
PR-003	Castellana 33	Punto de recarga elé...	Chamberí	Paseo Castellana 33	40,4319570	-3,6891439	441.547,43	4.475.
PR-004	Castellana 52	Punto de recarga elé...	Salamanca	Paseo Castellana 52	40,4355225	-3,6879659	441.650,00	4.476.
PR-005	Castellana 106	Punto de recarga elé...	Chamberí	Paseo Castellana 106	40,4454247	-3,6907790	441.420,49	4.477.
PR-006	Castellana 160	Punto de recarga elé...	Chamartín	Paseo Castellana 160	40,4584405	-3,6896385	441.528,45	4.478.
PR-007	Castellana 259	Punto de recarga elé...	Fuencarral-El Pardo	Paseo Castellana 259	40,4781350	-3,6864540	441.815,98	4.481.
PR-008	Castello 105	Punto de recarga elé...	Salamanca	Calle Castello 105	40,4349977	-3,6807567	442.261,15	4.476.
PR-009	Cerro de la plata 4	Punto de recarga elé...	Retiro	Calle Cerro de la Plata 4	40,3999038	-3,6752181	442.701,58	4.472.

Tableau es inteligente y nos ofrece la opción de marcar su Intérprete de datos

Ahora vemos el dataset que acabamos de cargar, y el anterior con el que hemos estado trabajando

No olvidaros de comprobar que vuestras variables están en el formato correcto

## Editar la relación entre los Datasets

Elegimos cuál es nuestro Dataset principal y cuál es el secundario

Barra de Herramientas > Datos > Editar relaciones

Elegimos la relación entre las variables de los datasets. Tableau es inteligente y detecta estas relaciones de forma automática. En nuestro ejemplo, el campo "Distrito" está presente en ambos datasets

# ¿Qué vamos a ver hoy?

1. Tableau. Interfaz Gráfica
2. Tableau. Arquitectura
3. Tableau para DS
4. Carto

1. Introducción
2. Conexión a los datos
3. Hojas de Trabajo
4. Gráficos predeterminados
5. Cálculos
6. Detección de valores atípicos
7. Combinación de datos
8. Eje Doble
9. Dashboards

# Eje Doble



Esta funcionalidad **nos permite comparar varias medidas** en dos ejes independientes, que se muestran como dos capas una encima de otra.



Son útiles cuando tenemos dos medidas en diferentes escalas. Equivaldría a la funcionalidad de **Excel** de pintar un gráfico en el **eje secundario**.



Para utilizarlo tenemos que disponer de **dos medidas** en el estante de filas o columnas, y en una de ellas seleccionar **clic derecho > Eje doble**. Así de sencillo

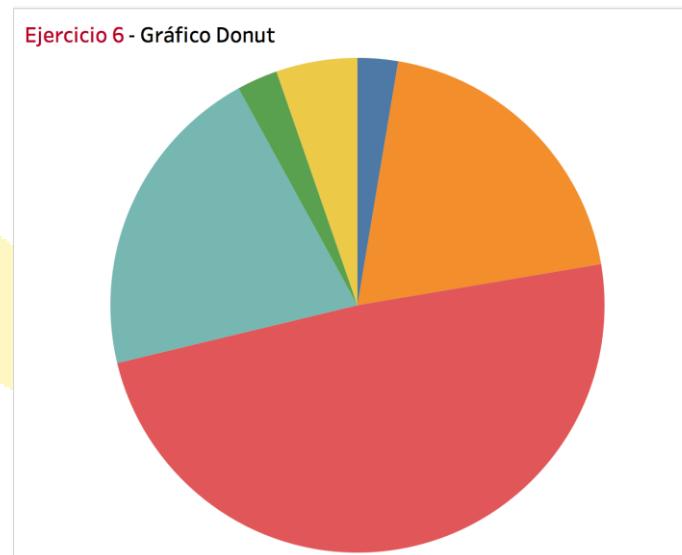
# Ejercicio



## Eje Doble. Gráfico Donut

### 1) Hacer un gráfico circular

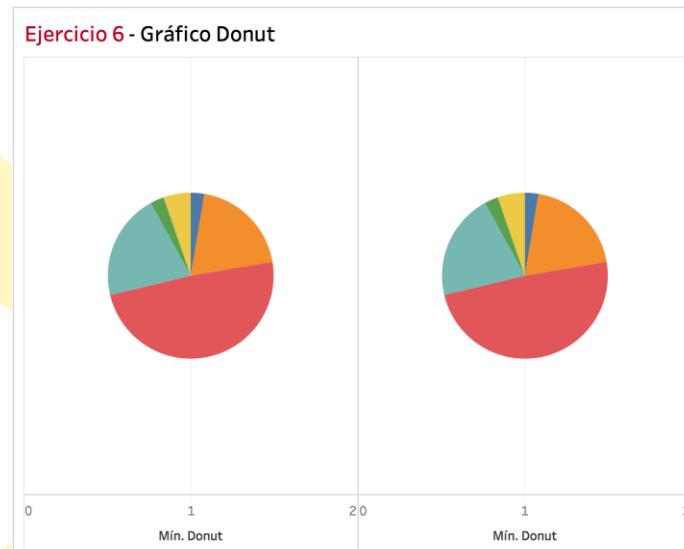
Es necesario añadir a nuestra visualización al menos una medida y una dimensión para poder elegir el gráfico circular en la tarjeta “Muéstrame”.



# Eje Doble. Gráfico Donut

## 2) Crear un eje de referencia

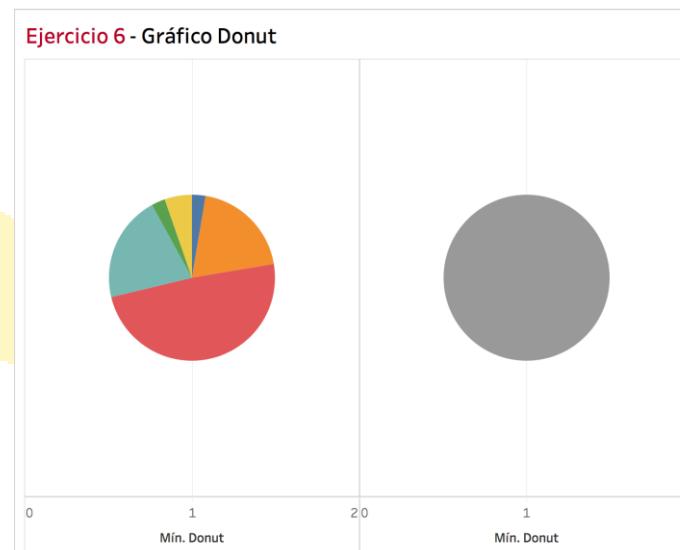
Como los gráficos circulares no tienen ejes, tenemos que hacer un pequeño truco para que funcione. Necesitamos crear un campo calculado que tome únicamente el valor 1. Una vez creado, arrastraremos este nuevo campo dos veces al estante columnas, agregado como Mínimo.



## Eje Doble. Gráfico Donut

### 3) Crear el agujero del anillo

Para ello, es necesario quitar las medidas y dimensiones añadidas al gráfico de la derecha, de modo que nos quede únicamente un círculo.



# Eje Doble. Gráfico Donut

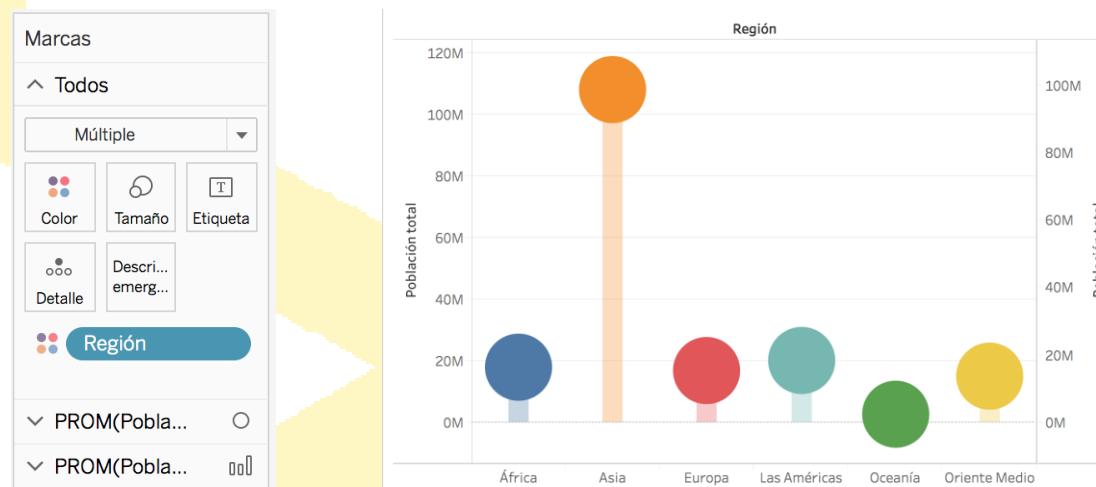
## 4) Utilizar el eje doble

Forzamos a que la visualización muestre el círculo “plano” en el eje doble. Después, cambiamos el color del círculo plano y su tamaño y obtenemos como resultado un gráfico Donut.

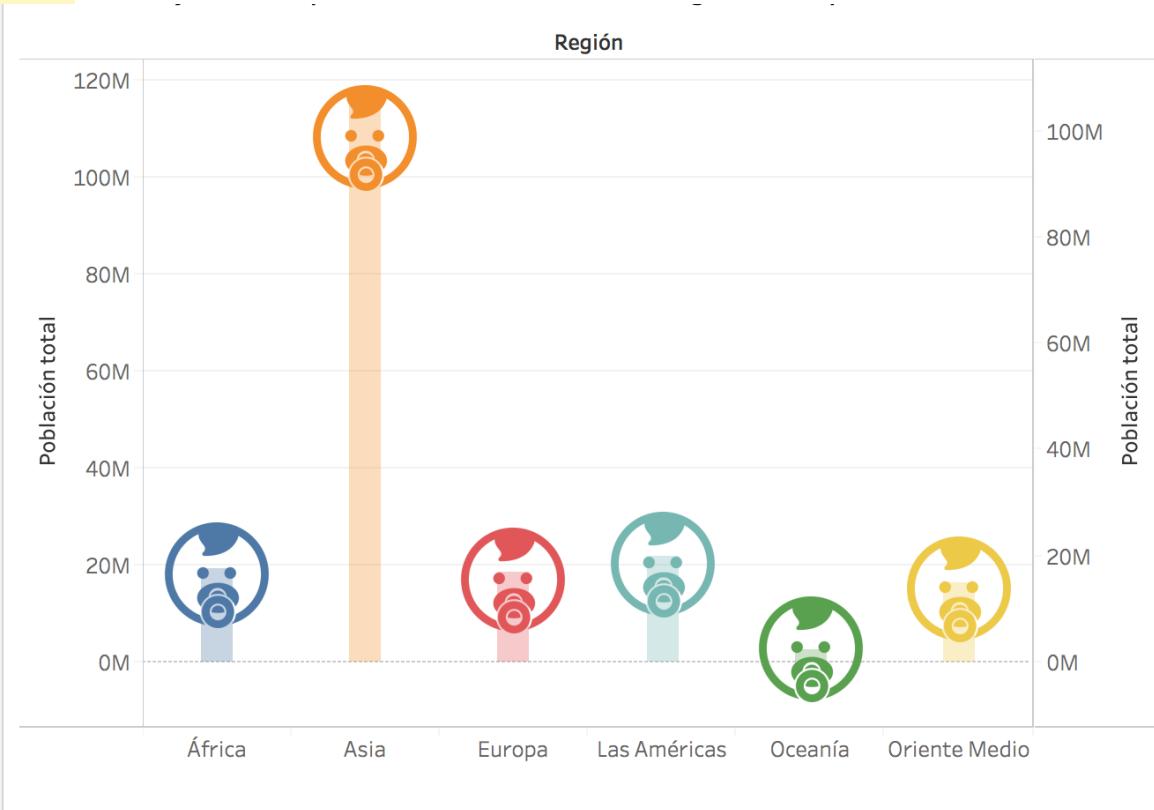


## Eje Doble. Gráfico “Piruleta”

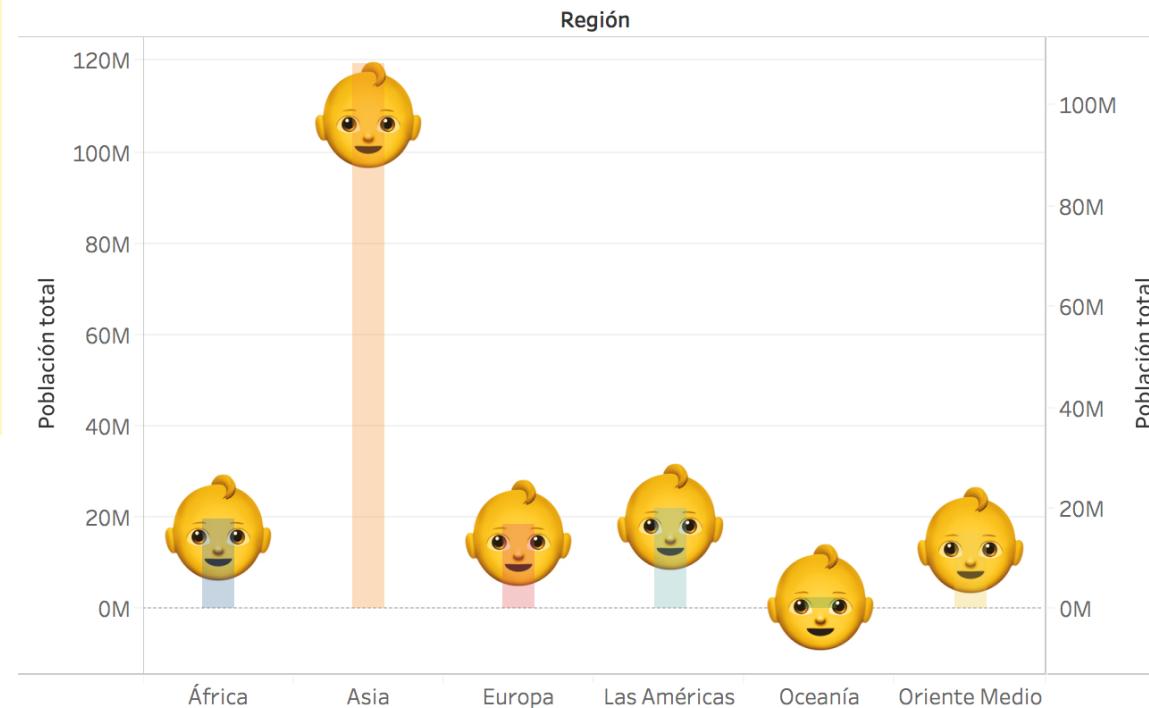
- 1) Seleccionar una medida y una dimensión.
- 2) “Duplicar” la misma medida en el estante filas y mostrarla en el eje doble.
- 3) Cambiar la forma de visualizar de las dos medidas: una como línea y la otra como círculo.



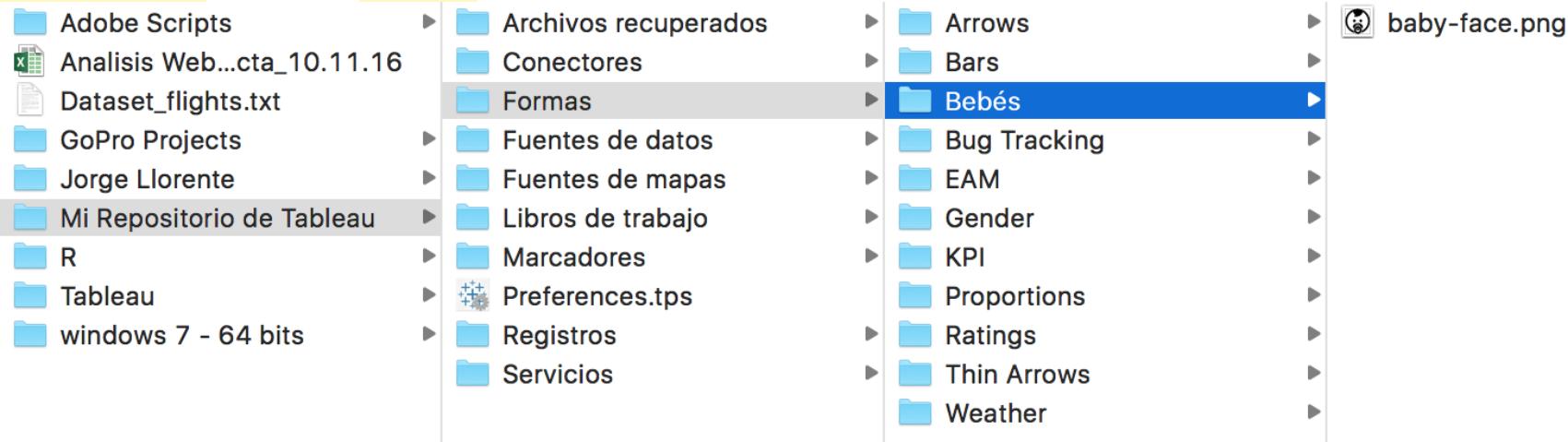
## Eje Doble. Gráfico de Piruleta con Formas Personalizadas



## Eje Doble. Gráfico de Piruleta con Formas Personalizadas



## Eje Doble. Gráfico de Piruleta con Formas Personalizadas. ¿Cómo?

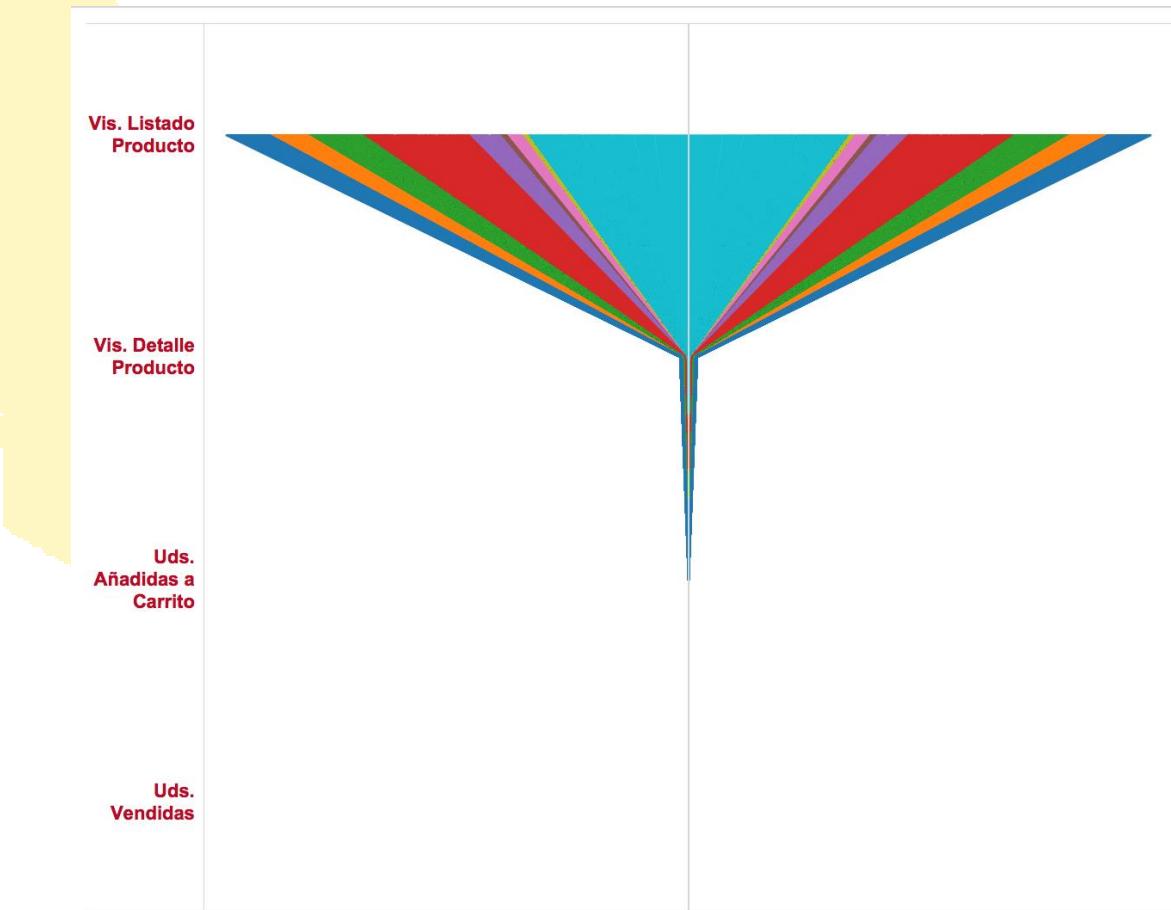


- 1) Debéis seguir la ruta Mi Repositorio de Tableau / Formas
- 2) Debéis crear una carpeta que contenga el grupo de iconos que queréis introducir
- 3) Pegáis el ícono dentro de la carpeta. Podéis obtener ideas de Flaticon.

# Ejercicio



## Otros gráficos. Embudo



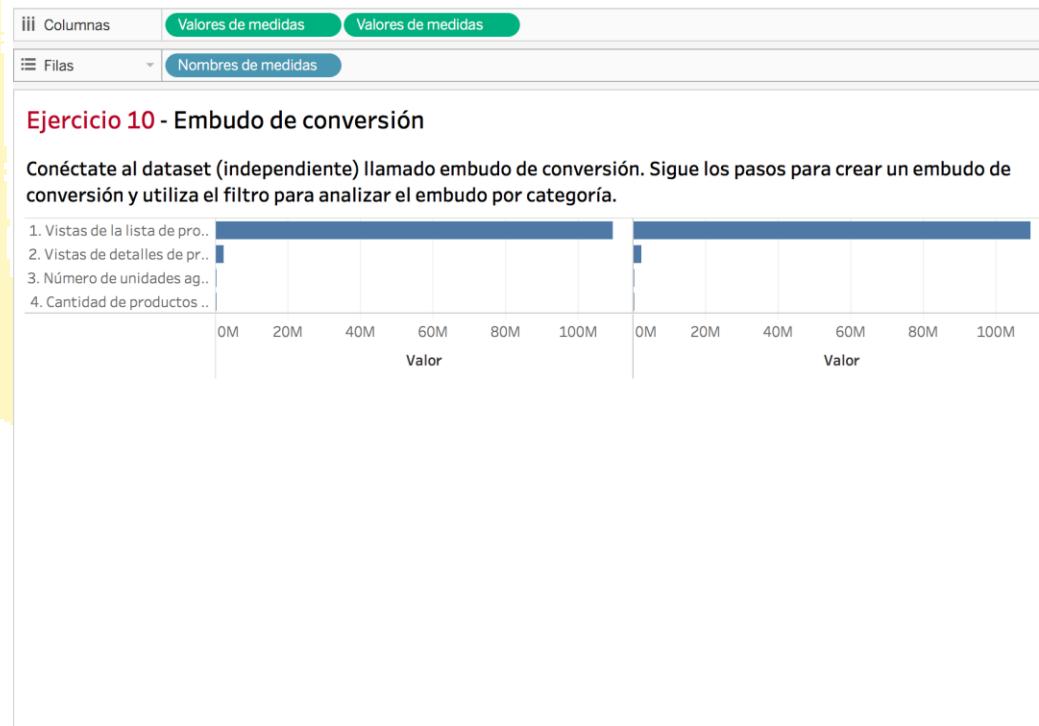
## Otros gráficos. Embudo

1) Llevamos dos veces la medida “**Valores de medidas**” al estante columnas. Veréis que en la tarjeta marcas se añaden todas las medidas, con lo cual seleccionamos únicamente las que nos interesan para confeccionar nuestro embudo y las colocamos en el orden correcto.



## Otros gráficos. Embudo

2) Llevamos la dimensión “Nombre de medidas” al estante filas para que se muestren los distintos pasos de nuestro proceso de compra.

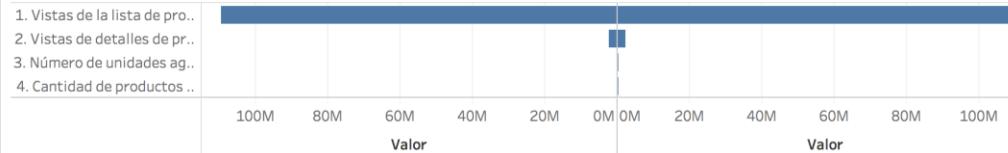


# Otros gráficos. Embudo

3) Editamos el eje de la izquierda para que se muestre de forma inversa al eje de la derecha (**pista**: botón derecho > editar eje)

## Ejercicio 10 - Embudo de conversión

Conéctate al dataset (independiente) llamado embudo de conversión. Sigue los pasos para crear un embudo de conversión y utiliza el filtro para analizar el embudo por categoría.

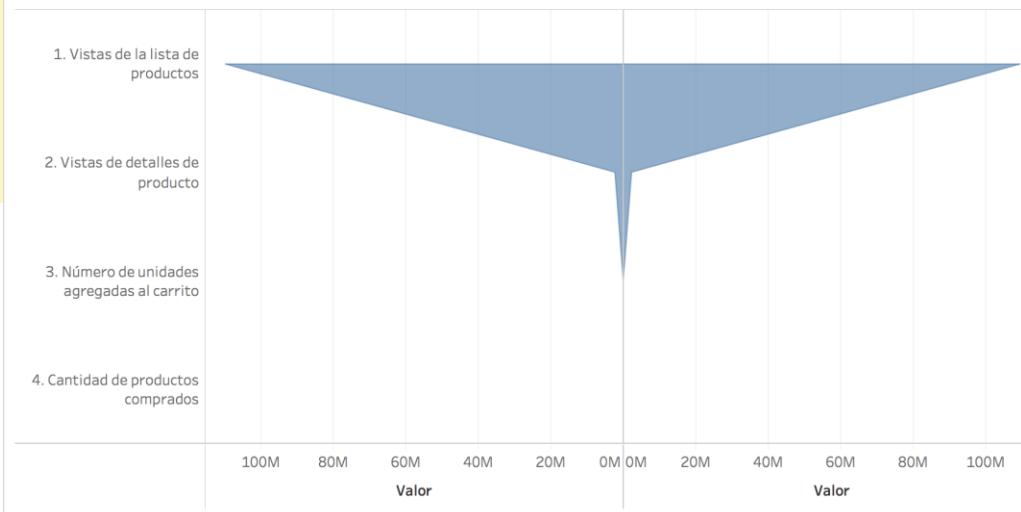


# Otros gráficos. Embudo

4) Cambiamos la forma de visualización para que nos muestre áreas en lugar de líneas. Voilà ☺

## Ejercicio 10 - Embudo de conversión

Conéctate al dataset (independiente) llamado embudo de conversión. Sigue los pasos para crear un embudo de conversión y utiliza el filtro para analizar el embudo por categoría.



# ¿Qué vamos a ver hoy?

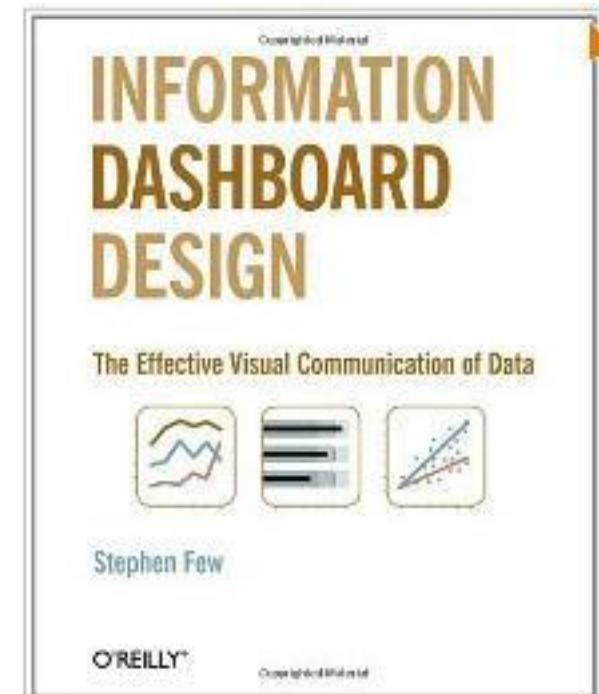
1. Tableau. Interfaz Gráfica
2. Tableau. Arquitectura
3. Tableau para DS
4. Carto

1. Introducción
2. Conexión a los datos
3. Hojas de Trabajo
4. Gráficos predeterminados
5. Cálculos
6. Detección de valores atípicos
7. Combinación de datos
8. Eje Doble
9. Dashboards

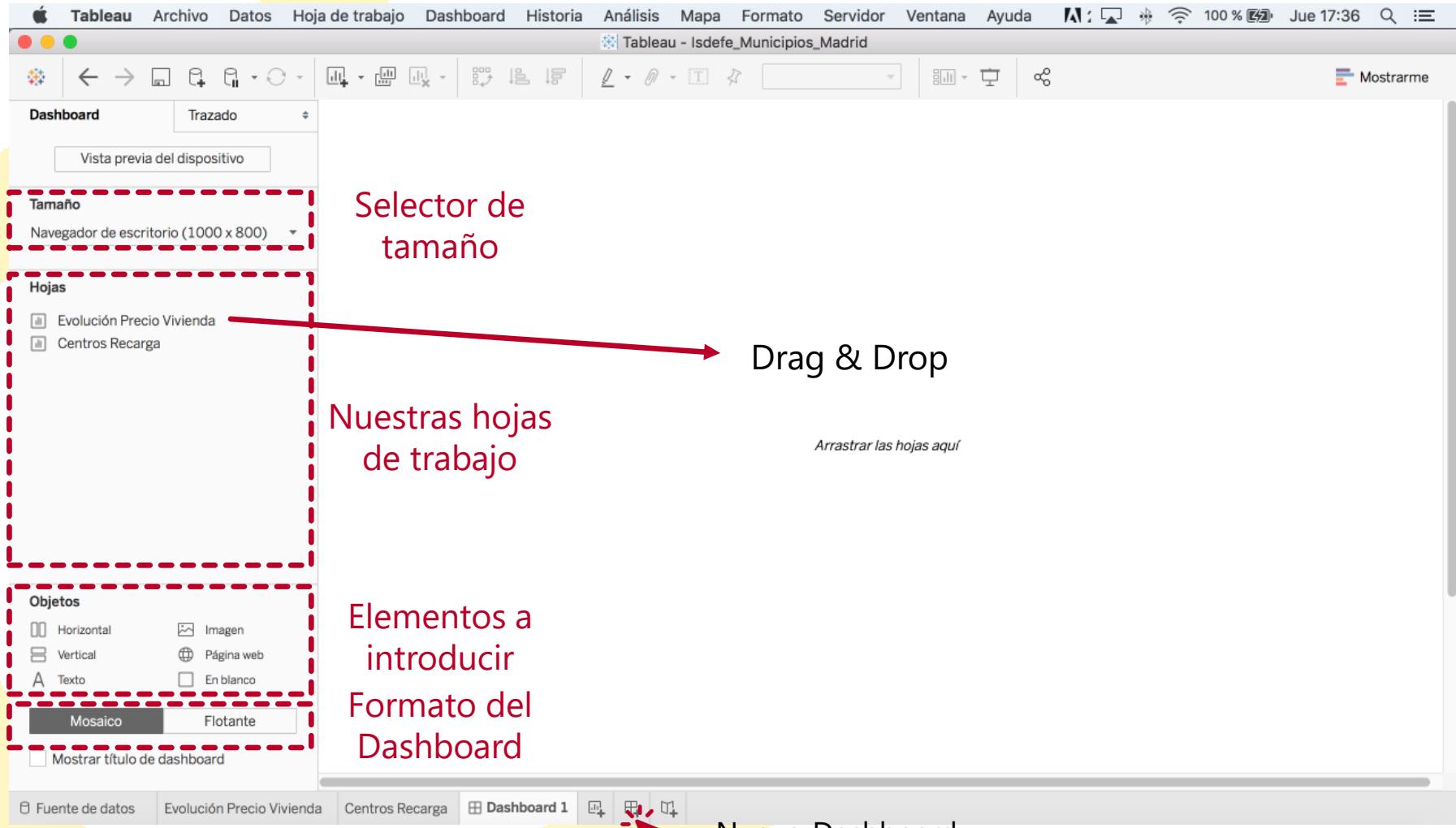
# Dashboards

*"Un dashboard es un panel que contiene los elementos más importantes de información necesarios para conseguir uno o más objetivos; consolidado y creado como una única pantalla donde toda la información puede ser revisada de un vistazo".*

*Stephen Few, 2004*



# Dashboards en Tableau



The screenshot shows the Tableau desktop application. On the left, there's a sidebar with three main sections: 'Tamaño' (Size) set to 'Navegador de escritorio (1000 x 800)', 'Hojas' (Sheets) containing 'Evolución Precio Vivienda' and 'Centros Recarga', and 'Objetos' (Objects) with options for 'Horizontal', 'Vertical', 'Text', 'Imagen' (Image), 'Página web' (Web Page), 'En blanco' (Blank), 'Mosaico' (Mosaic), and 'Flotante' (Floating). At the bottom of this sidebar is a checkbox for 'Mostrar título de dashboard' (Show dashboard title). The main workspace is titled 'Tableau - Isdefe\_Municipios\_Madrid'. It features a toolbar at the top with various icons for file operations, and a large central area labeled 'Arrastrar las hojas aquí' (Drag sheets here) where a red arrow points from the 'Hojas' section of the sidebar. In the bottom right corner of the workspace, there's a button with a plus sign and a red circle with a white exclamation mark, which is highlighted with a yellow callout bubble containing the text 'Nuevo Dashboard' (New Dashboard).

Selector de tamaño

Nuestras hojas de trabajo

Elementos a introducir

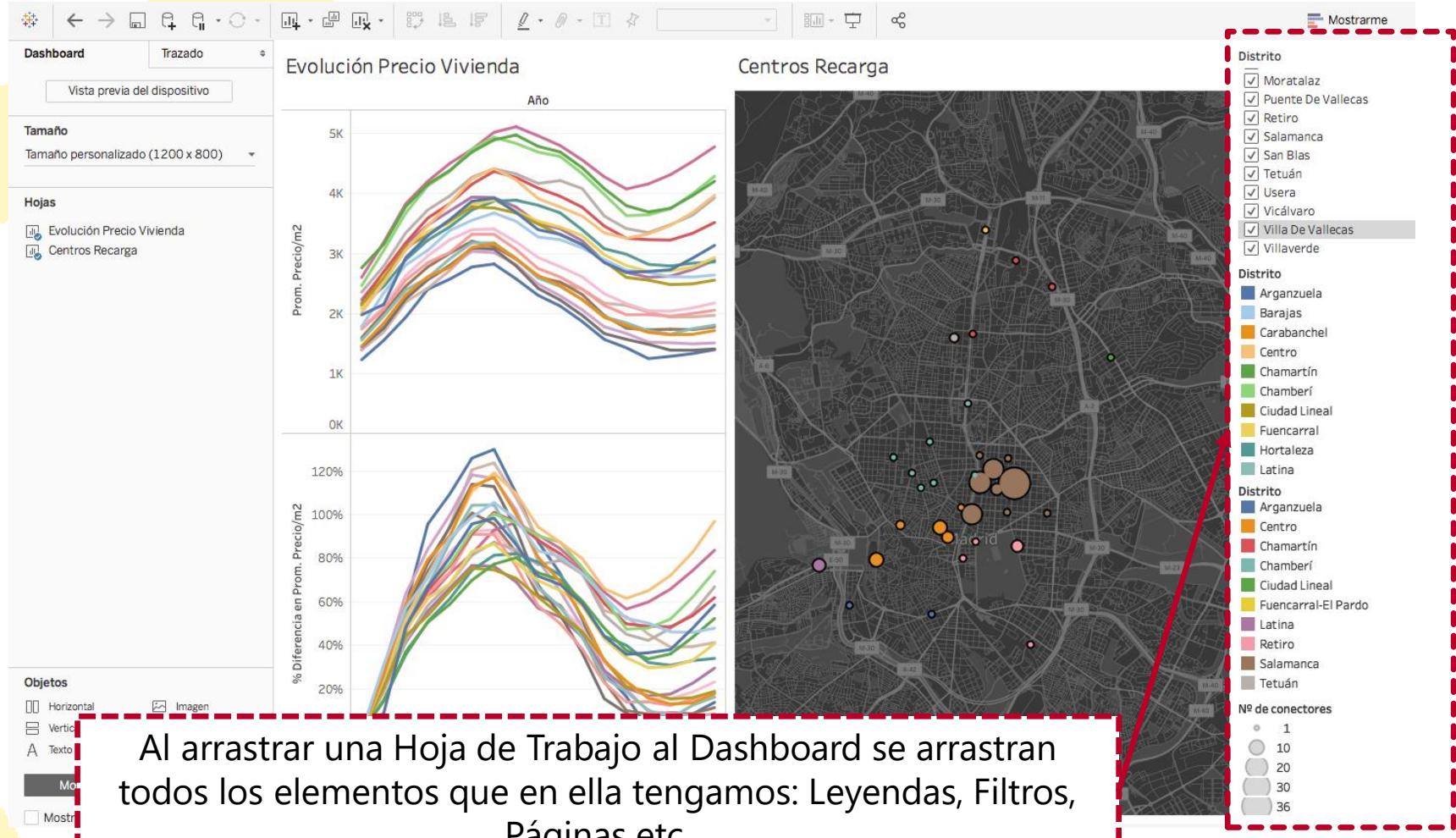
Formato del Dashboard

Drag & Drop

Arrastrar las hojas aquí

Nuevo Dashboard

# Dashboards en Tableau



The screenshot shows a Tableau dashboard titled "Evolución Precio Vivienda". It contains two line charts and a map of Madrid. The left chart, "Evolución Precio Vivienda", displays the average price per square meter over time for various districts. The right chart, "Centros Recarga", shows the locations of charging stations across the city. A legend on the right lists districts and their corresponding colors. A yellow callout box highlights the sidebar, and a red dashed box highlights the map area.

Al arrastrar una Hoja de Trabajo al Dashboard se arrastran todos los elementos que en ella tengamos: Leyendas, Filtros, Páginas etc.

**Dashboard** Trazado

Vista previa del dispositivo

Tamaño

Tamaño personalizado (1200 x 800)

Hojas

- Evolución Precio Vivienda
- Centros Recarga

Objetos

- Horizontal
- Vertical
- A Texto
- Mostrar
- Mostrar

**Evolución Precio Vivienda**

Prom. Precio/m<sup>2</sup>

Año

% Diferencia en Prom. Precio/m<sup>2</sup>

**Centros Recarga**

Mostrarme

**Distrito**

- Moratalaz
- Puente De Vallecas
- Retiro
- Salamanca
- San Blas
- Tetuán
- Usera
- Vicálvaro
- Villa De Vallecas
- Villaverde

**Distrito**

- Arganzuela
- Barajas
- Carabanchel
- Centro
- Chamartín
- Chamberí
- Ciudad Lineal
- Fuencarral
- Hortaleza
- Latina

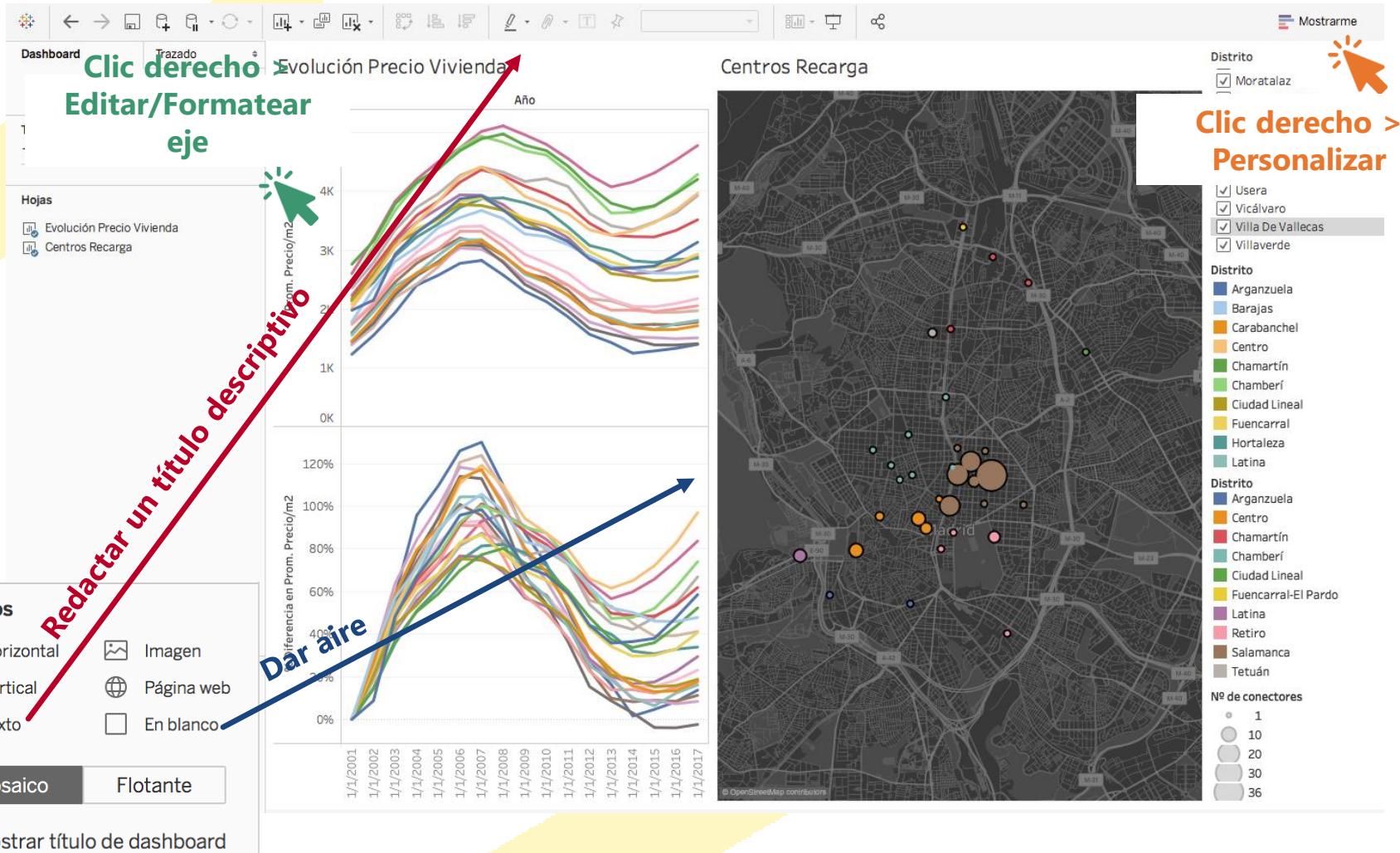
**Distrito**

- Arganzuela
- Centro
- Chamartín
- Chamberí
- Ciudad Lineal
- Fuencarral-El Pardo
- Latina
- Retiro
- Salamanca
- Tetuán

**Nº de conectores**

- 1
- 10
- 20
- 30
- 36

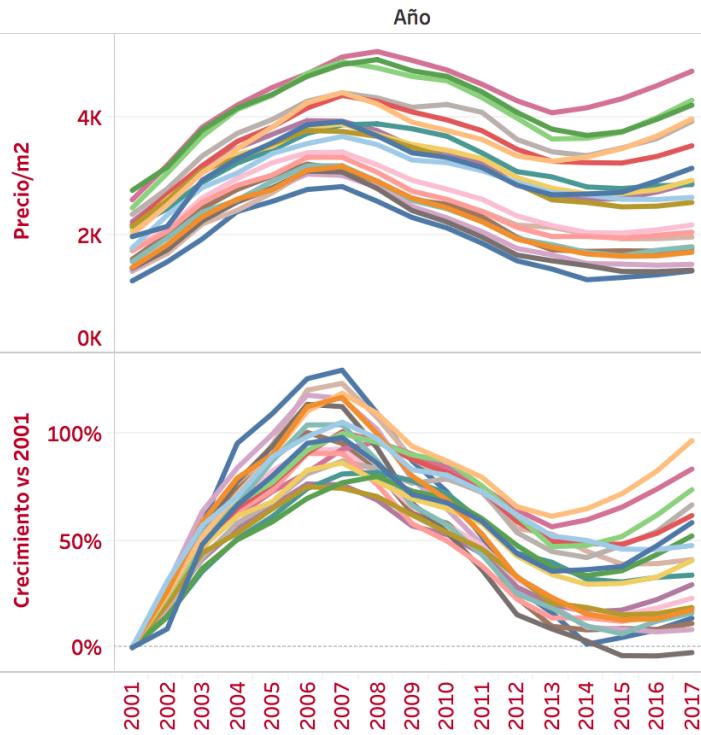
# ¿Qué puedo hacer para optimizarlo?



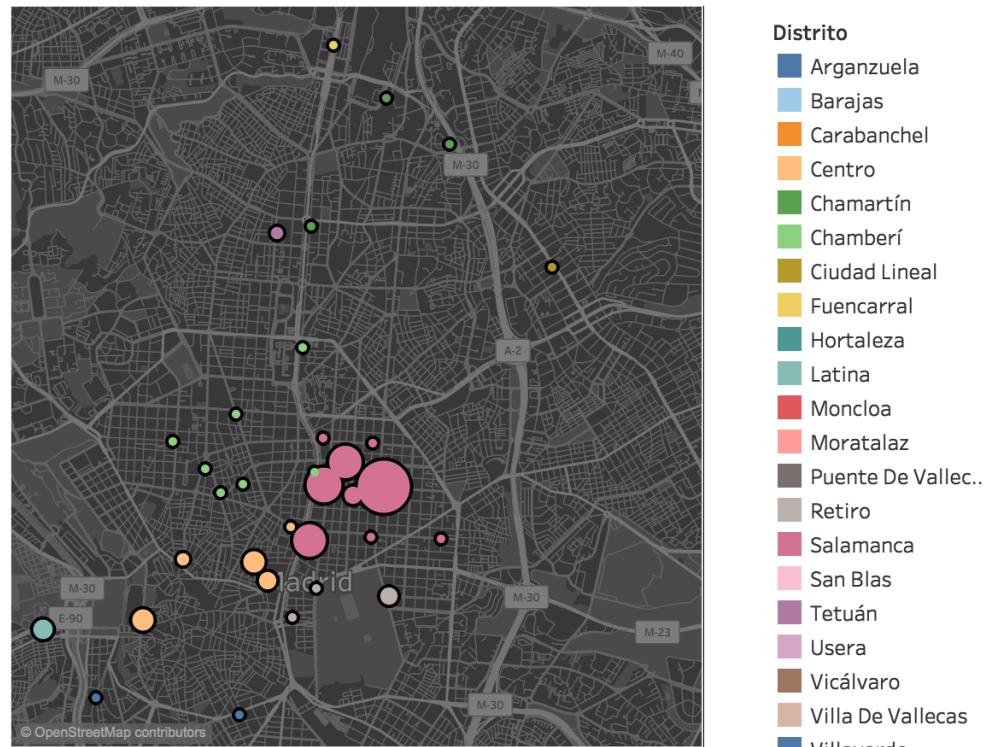
# Mucho mejor, ¿no?

¿Donde comprarme una casa dadas mis condiciones actuales?

Evolución Precio Vivienda



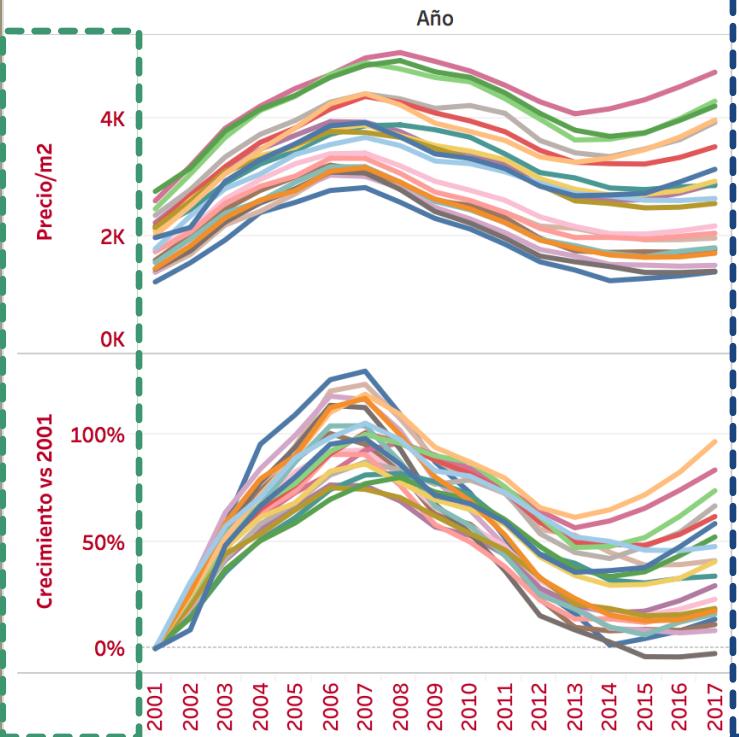
Centros Recarga



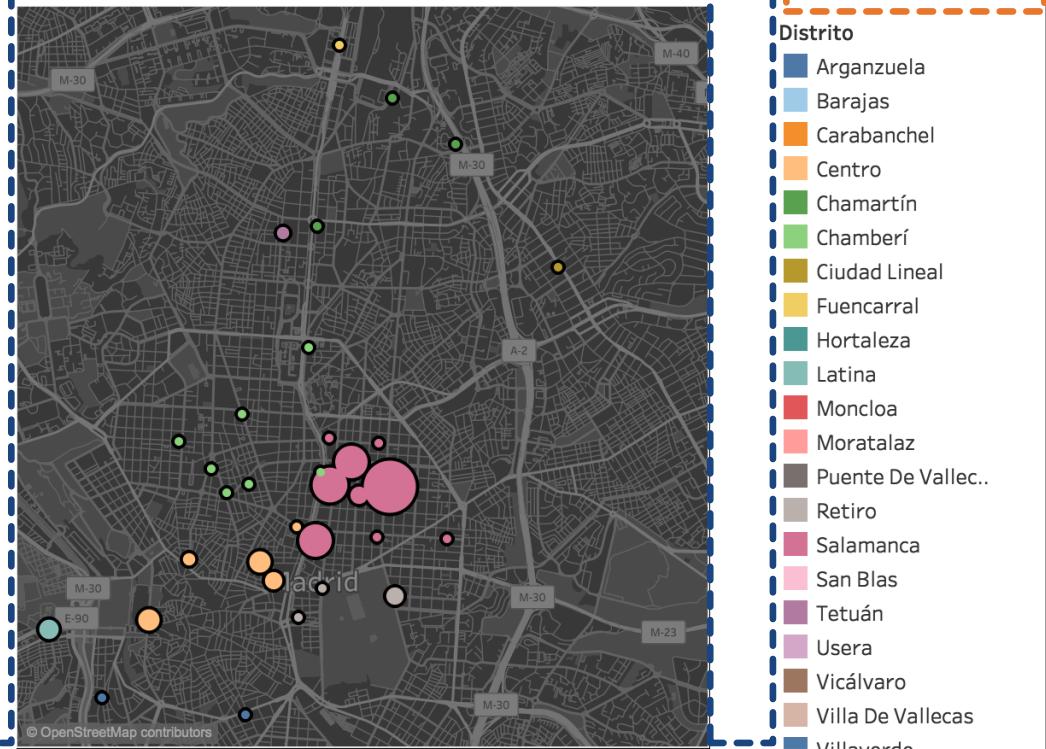
# Recapitulando

¿Donde comprarme una casa dadas mis condiciones actuales?

Evolución Precio Vivienda

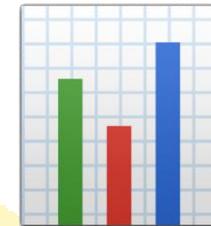


Centros Recarga



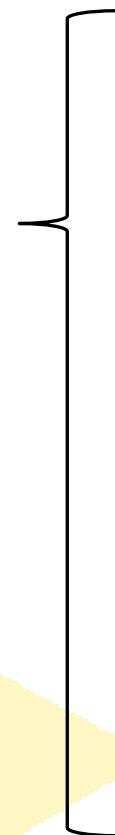
## Vamos a practicar

Abrid el fichero “Enunciado\_Ejercicios” y haced todos los demás ejercicios. Sentiros libres de “dejar volar” vuestra creatividad.



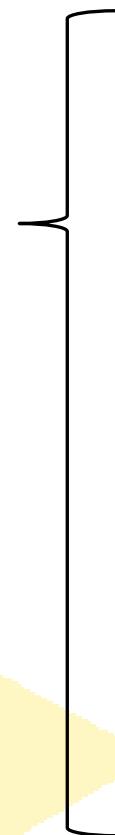
# ¿Qué vamos a ver?

- 
- A large, semi-transparent yellow chevron shape points from the bottom left towards the center of the slide, containing the following list:
1. Tableau. Interfaz Gráfica
  2. Tableau. Arquitectura
  3. Tableau para DS
  4. Carto

- 
- A black curly brace on the right side groups the following content:
1. Arquitectura
    - Live Query Engine & In Memory Engine
    - Conectividad
  2. Estrategias de presentación
    - Temperatura de los datos (Cold, warm, hot)
    - Qué hacer y qué no...

## ¿Qué vamos a ver?

- 
- A large, semi-transparent yellow chevron shape points from the bottom left towards the center of the slide, containing the following list:
1. Tableau. Interfaz Gráfica
  2. Tableau. Arquitectura
  3. Tableau para DS
  4. Carto

- 
- A black curly brace on the right side groups the following content:
1. Arquitectura
    - Live Query Engine & In Memory Engine
    - Conectividad
  2. Estrategias de presentación
    - Temperatura de los datos (Cold, warm, hot)
    - Qué hacer y qué no...

## Arquitectura Híbrida de Datos

- Diseñada para aportar adaptabilidad y flexibilidad
- 'Transparente' al usuario
- Distintas soluciones para distintos problemas:
  - Live Query Engine
  - In-Memory Engine
  - Una 'mezcla' de ambos

# Arquitectura Híbrida de Datos (II)

- **Live Query Engine**

- Conexión directa a los data stores
- Libera carga en el servidor, delegándola en los orígenes de datos (push-down workload)
- Los datos permanecen en la fuente

# Arquitectura Híbrida de Datos (II)

- **In-Memory Engine**

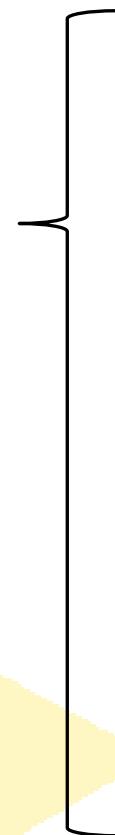
- Carga los datos en el Tableau Server (push-up workload)
- Permite agregar los datos de origen de forma previa a su carga
- Los datos a procesar no se ven afectados por cambios en el origen

# Arquitectura Híbrida de Datos (III)

Live Query	PRESENTACIÓN	In-Memory Query
Cualquier volumen, cualquier granularidad	<b>VOLUMEN DE LOS DATOS</b>	Inviable con grandes volumenes
Inmediata	<b>DISPONIBILIDAD DE LOS DATOS</b>	TDE's -> Retraso
Conexión directa -> Simple	<b>SIMPLICIDAD</b>	Refresco y gestión de TDE's -> Complejo
Limitado por el origen de datos, y red	<b>RENDIMIENTO VELOCIDAD</b>	Alto rendimiento al procesar en memoria
Consultas consecutivas pueden dar distintos resultados	<b>ESTABILIDAD DE LOS DATOS</b>	No afectado por cambios en origen

# ¿Qué vamos a ver?

- 
- A large, semi-transparent yellow chevron shape points from the bottom left towards the center of the slide, containing the following list:
1. Tableau. Interfaz Gráfica
  2. Tableau. Arquitectura
  3. Tableau para DS
  4. Carto

- 
- A black curly brace on the right side groups the following list into two main categories:
1. Arquitectura
    - Live Query Engine & In Memory Engine
    - Conectividad
  2. Estrategias de presentación
    - Temperatura de los datos (Cold, warm, hot)
    - Qué hacer y qué no...

# Estrategias de Presentación

- Distintos escenarios, según el origen de los datos:



COLD

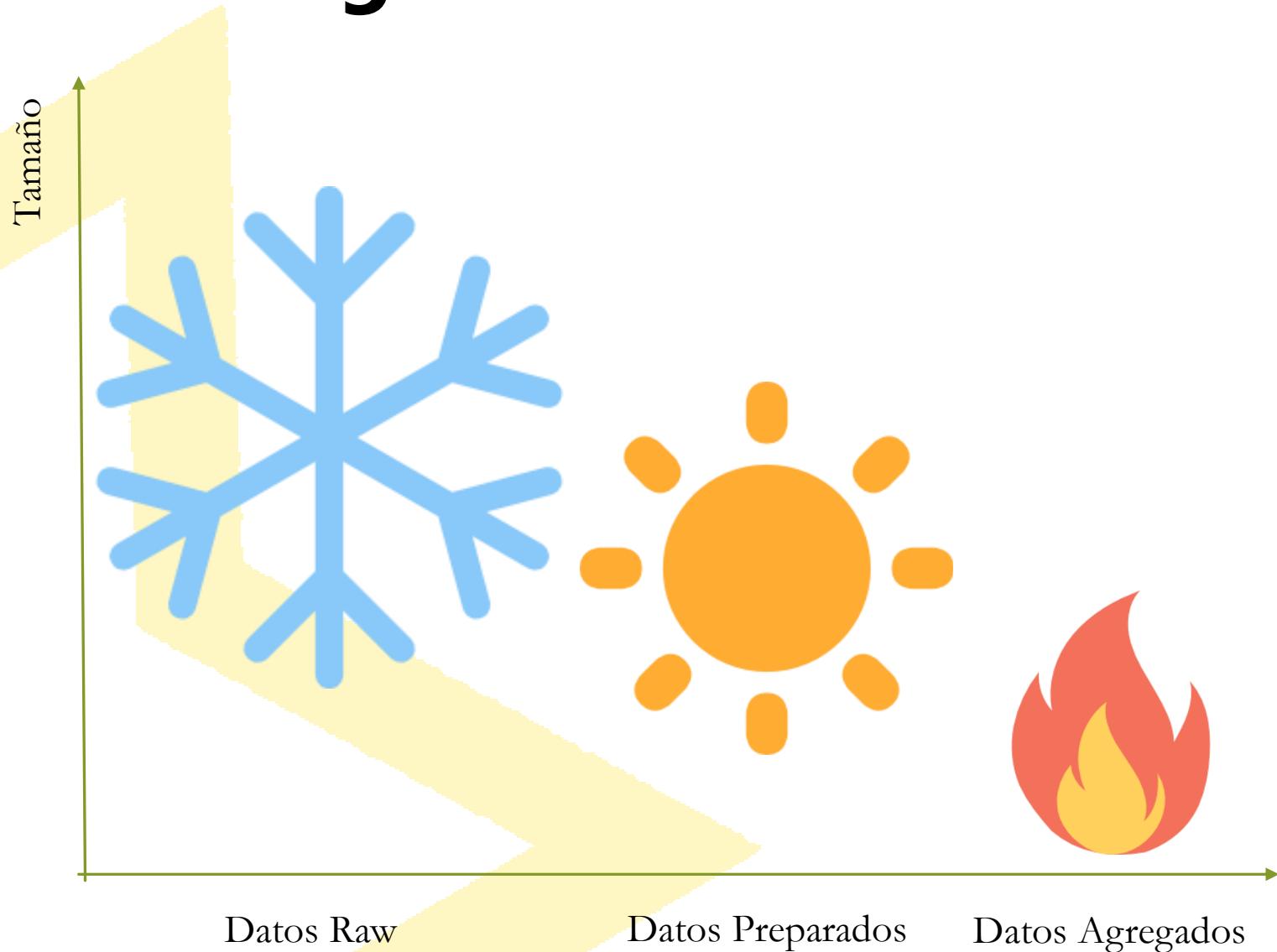


WARM

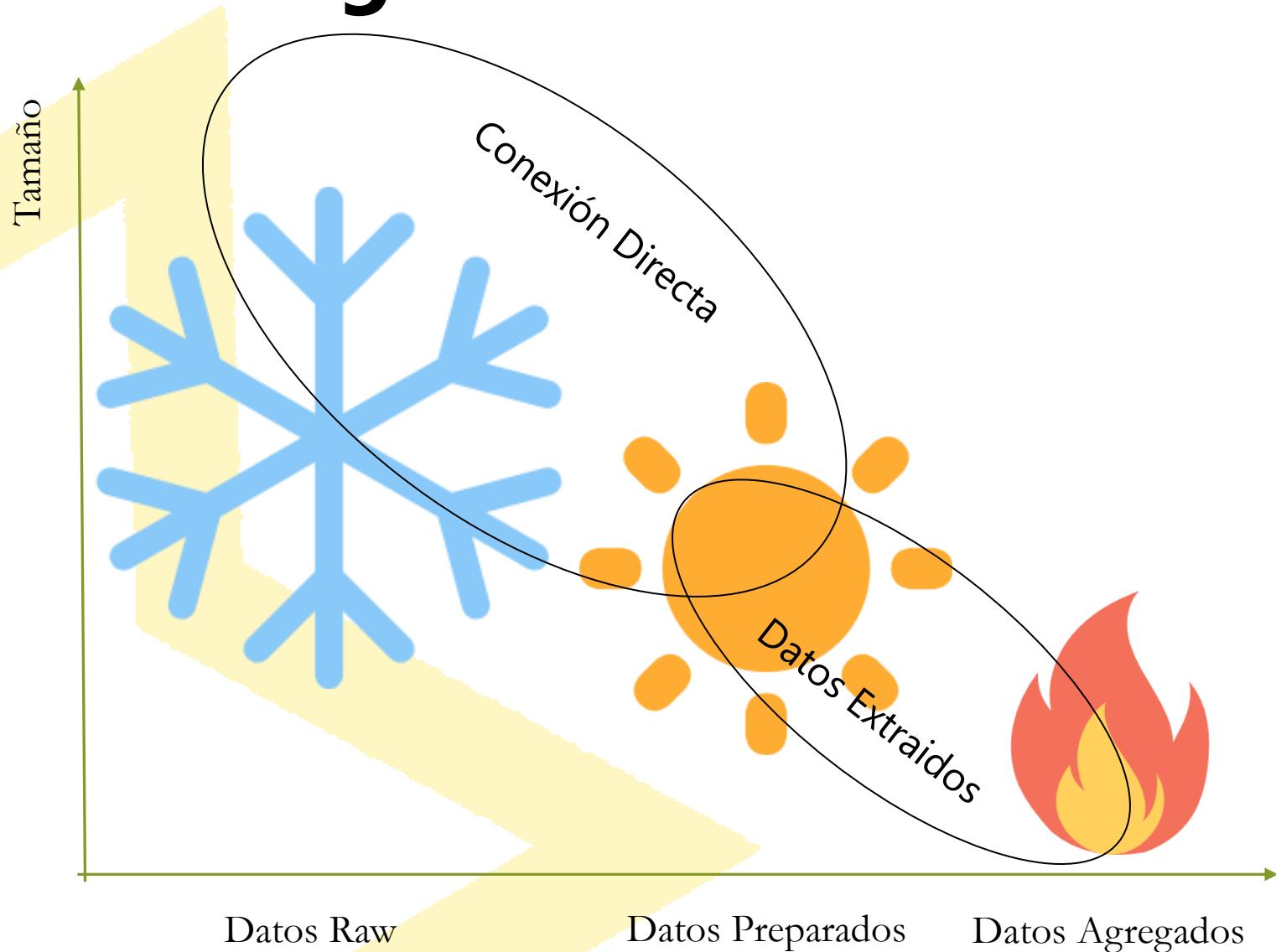


HOT

# Estrategia de Presentación II



# Estrategia de Presentación II



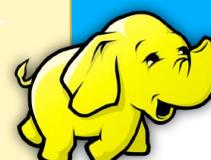
## Cold Data

- **CASOS DE USO**

- *"Almacena todo, analiza algo"*
- Exploración y minería de datos



HDFS

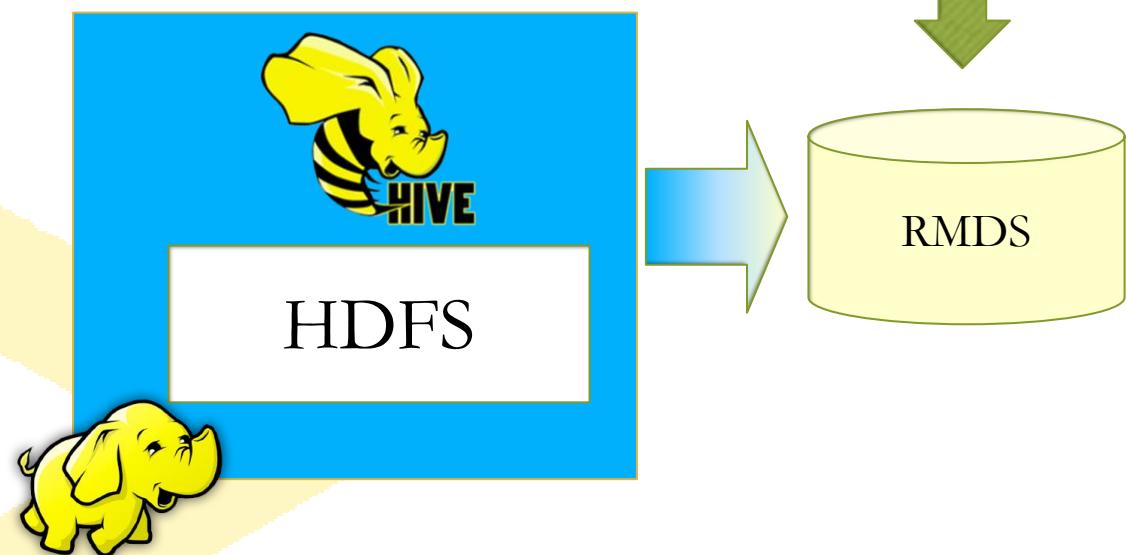


## Warm Data



- **CASOS DE USO**

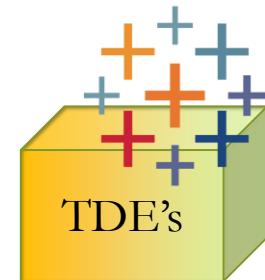
- Datos e Informes migrados desde Hadoop, o históricos
- Datos analizados periódicamente



## Hot Data

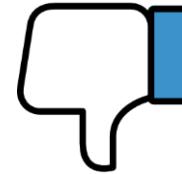


- **CASOS DE USO**
  - Aplicaciones BI
  - Datos maestros
  - Datos con requisitos específicos  
(aislamiento del origen, inmediatez en la consulta, ...)



## Extracciones Vs Live Connections

- No siempre los TDE's son la mejor solución:

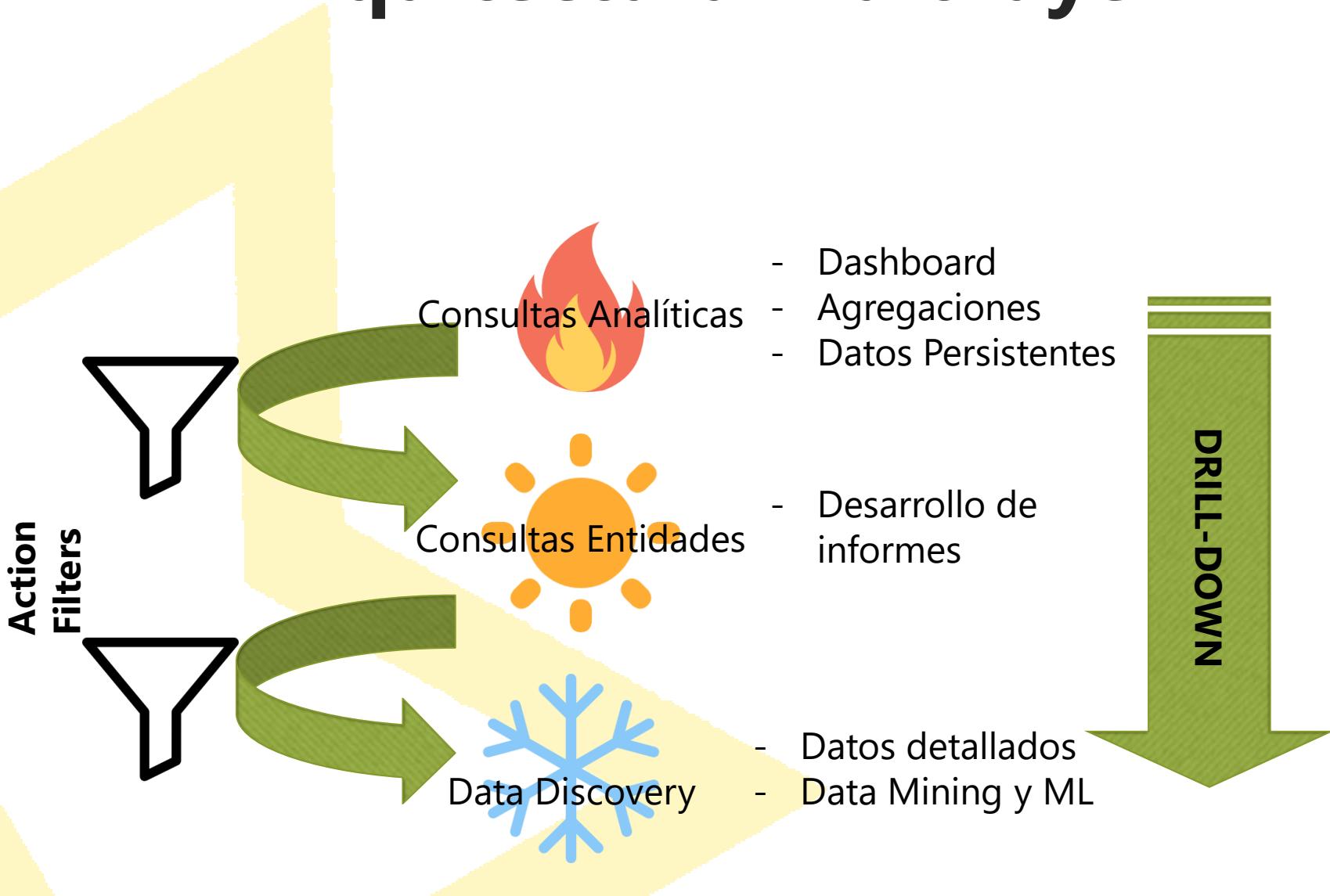


- Volúmenes de datos limitados
- Se requiere análisis en offline
- No se desea sobrecargar el cluster de Hadoop
- Motor de SQL-on-Hadoop potente
- Se requiere análisis en tiempo real

# ¿Cómo optimizamos los TDE's?

- Antes de la extracción:
  - Pre-agregando los datos
  - Realizando pre-computaciones y almacenamiento de los resultados
  - Evitando los JOIN innecesarios
  - Eliminando información innecesaria (hidden columns y subset filters),

# Arquitectura Multilayer



## ¿Qué vamos a ver?

- 
- 1. Tableau. Interfaz Gráfica
  - 2. Tableau. Arquitectura
  - 3. Tableau para DS
  - 4. Carto
- 1. Integración con R
  - 2. Integración con Python

## Integración con R

- De forma nativa, es posible desde Tableau (en licencias desktop y de servidor Tableau Server), ejecutar scripts desarrollados en R
- Para ello, solo es necesario que desde la instalación de Tableau se acceda a un servidor de R via Rserve.
- La integración se realiza mediante la definición de campos calculados que utilizan funciones específicas para scripting.

## Funciones de scripting

- Entre las funciones predefinidas, para la definición de campos calculados en Tableau, encontramos el conjunto de funciones **SCRIPT\_\***:
  - `SCRIPT_REAL`, `SCRIPT_STR`, `SCRIPT_INT` y `SCRIPT_BOOL`
- La diferencia principal entre ellas, reside en el tipo de dato que devuelve (según el nombre de la función)
- Estas funciones permiten invocar funciones de R, incluso con paso de parámetros, y devolver a Tableau el resultado de la evaluación.

## Limitaciones

- No se pueden exportar datos de Tableau a R de manera directa (solo a través de las funciones de SCRIPT\_\*).
- No se pueden exportar visualizaciones de R a Tableau
- No se pueden importar conjuntos de datos de R a Tableau. Solo se pueden obtener a través de las funciones mencionadas

## Instalación

- Descarga e instala R
- Descarga e instala Rserve (desde la consola de R:

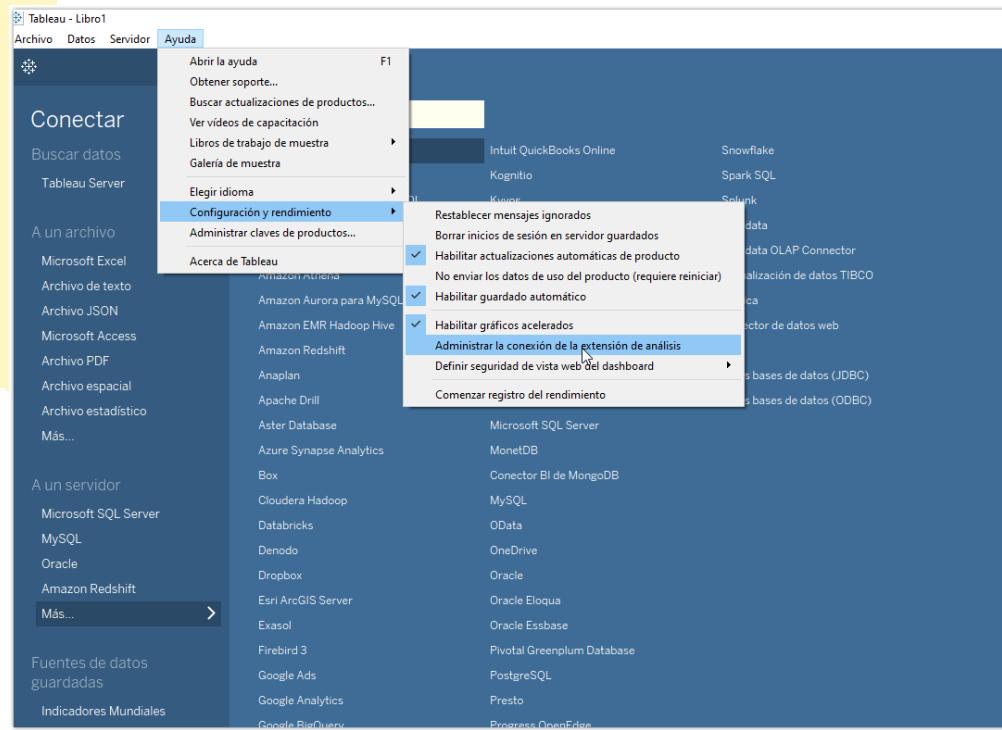
```
Install.packages("Rserve")
```

```
Library(Rserve)
```

```
Rserve()
```

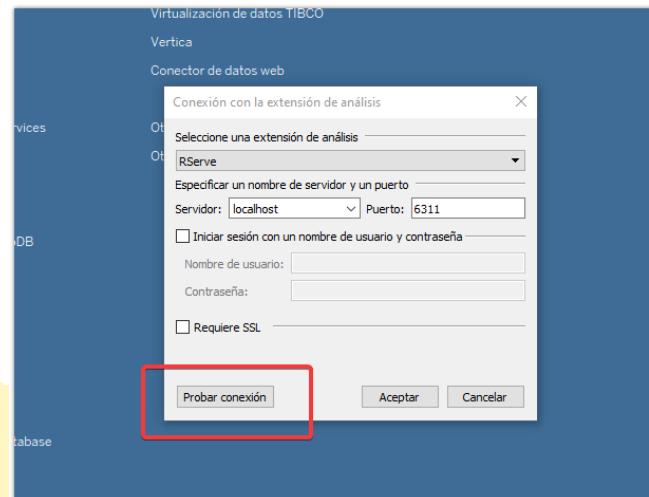
# Instalación II

- Conecte Tableau a Rerve
  - Desde ayuda, acceda a:



## Instalación III

- Seleccione “Rserve” como extensión de análisis, y especifique “localhost” como servidor (puerto 6311)



- Pruebe la conexión desde el botón habilitado al efecto (debería indicar que la conexión se estableció correctamente)

## Uso

- Una vez instalado, desde Tableau podremos invocar funciones de las librerías disponibles, incluyendo la lógica en forma de parámetro de la función SCRIPT\_\* correspondiente.
- Ejemplo:

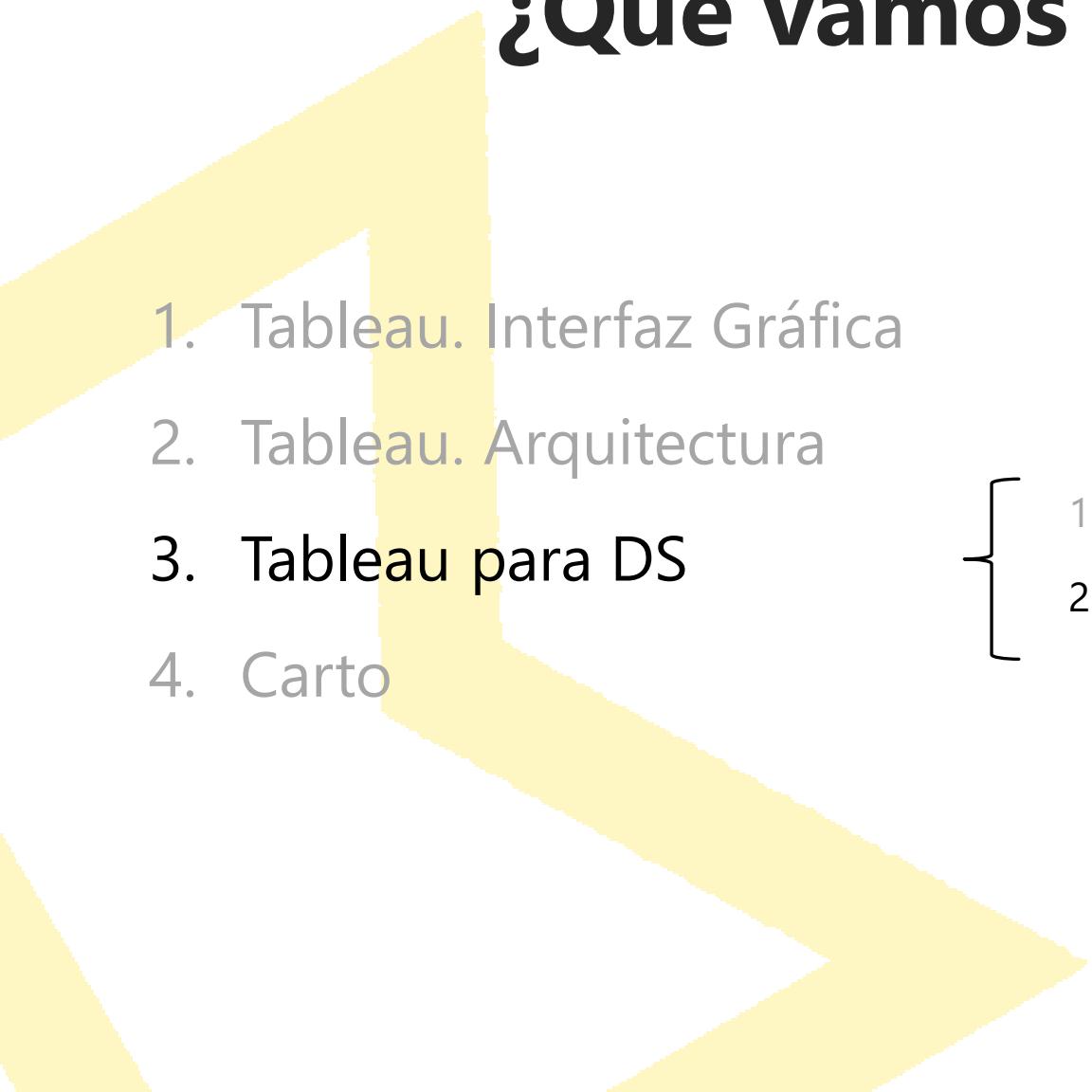
```
RSCRIPT_BOOL ( “(.arg1 + .arg2)/2”, MIN([Datos]),  
MAX([Datos]) )
```

- NOTA: Es posible pasar parámetros – por posición – a las funciones SCRIPT\_\* a través de las palabras reservadas .arg1, .arg2, ...

# Ejercicio



## ¿Qué vamos a ver?

- 
- 1. Tableau. Interfaz Gráfica
  - 2. Tableau. Arquitectura
  - 3. Tableau para DS
  - 4. Carto
- 1. Integración con R
  - 2. Integración con Python

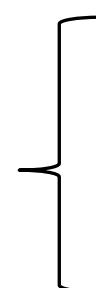
## Integración con Python

- De forma análoga, es posible invocar desde Tableau funciones en Python, o utilizar librerías como NumPy o Pandas.
- El procedimiento de configuración de Tableau es análogo al descrito para R, salvo que en este caso la conexión se habilitará a la extensión de análisis [TabPy](#), en lugar de a la de Rserve, la cual deberá instalarse previamente desde la consola de Python.

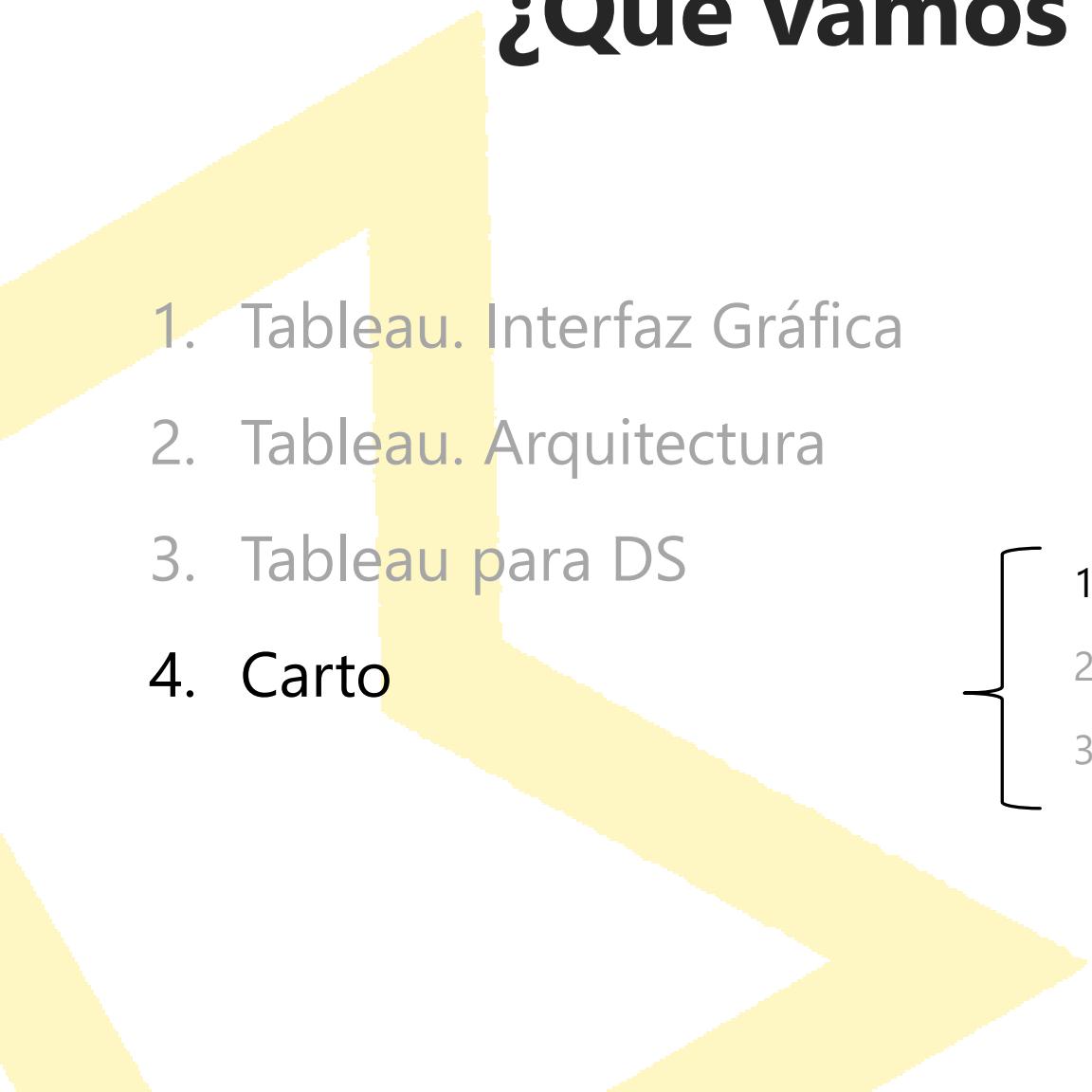
## Uso de parámetros

- Al igual que en el caso de la integración con R, al invocar las funciones SCRIPT\_\*, es posible pasar parámetros al script de Python invocado.
- En este caso, se utilizarán las palabras reservadas \_arg1, \_arg2, ... en lugar de .arg1, .arg2, ...

## ¿Qué vamos a ver?

- 
1. Tableau. Interfaz Gráfica
  2. Tableau. Arquitectura
  3. Tableau para DS
  4. Carto
- 
  1. Introducción
  2. Módulos
  3. CartoBuilder

## ¿Qué vamos a ver?

- 
- 1. Tableau. Interfaz Gráfica
  - 2. Tableau. Arquitectura
  - 3. Tableau para DS
  - 4. Carto
- 1. Introducción
  - 2. Módulos
  - 3. CartoBuilder

## Introducción

- Inicialmente CartoDB
- Desarrollado desde 2008 por startup española (<http://www.vizzuality.com>)
- Candidata a startup “unicornio” en 2017
- Proyecto opensource (<https://github.com/CartoDB/cartodb>)

## ¿Qué es y para qué sirve?

- Es una herramienta GIS opensource, con una enorme base de datos geoespaciales, y capacidad de visualización
- Permite el análisis de la información y su presentación web.
- Encaminada a la toma de decisiones basada en la contextualización geográfica (Location Intelligence)
- Modelo de negocio: Freemium

## Módulos

- Carto Builder
  - Aplicación web dirigida a la presentación visual de mapas y datos asociados
  - No necesita conocimientos de programación
- Carto Engine
  - Motor de ejecución, con API's de desarrollo y librerías .js
    - [CARTO.js](#)
    - [SQL API](#)
    - Data Services API

# Carto Builder



Add datasets  
Select your datasets

SEARCH CONNECT DATASET YOUR DATASETS DATA LIBRARY CREATE EMPTY MAP

You have not connected any datasets yet  
You can [connect your own dataset](#) or choose one from our Data Library.

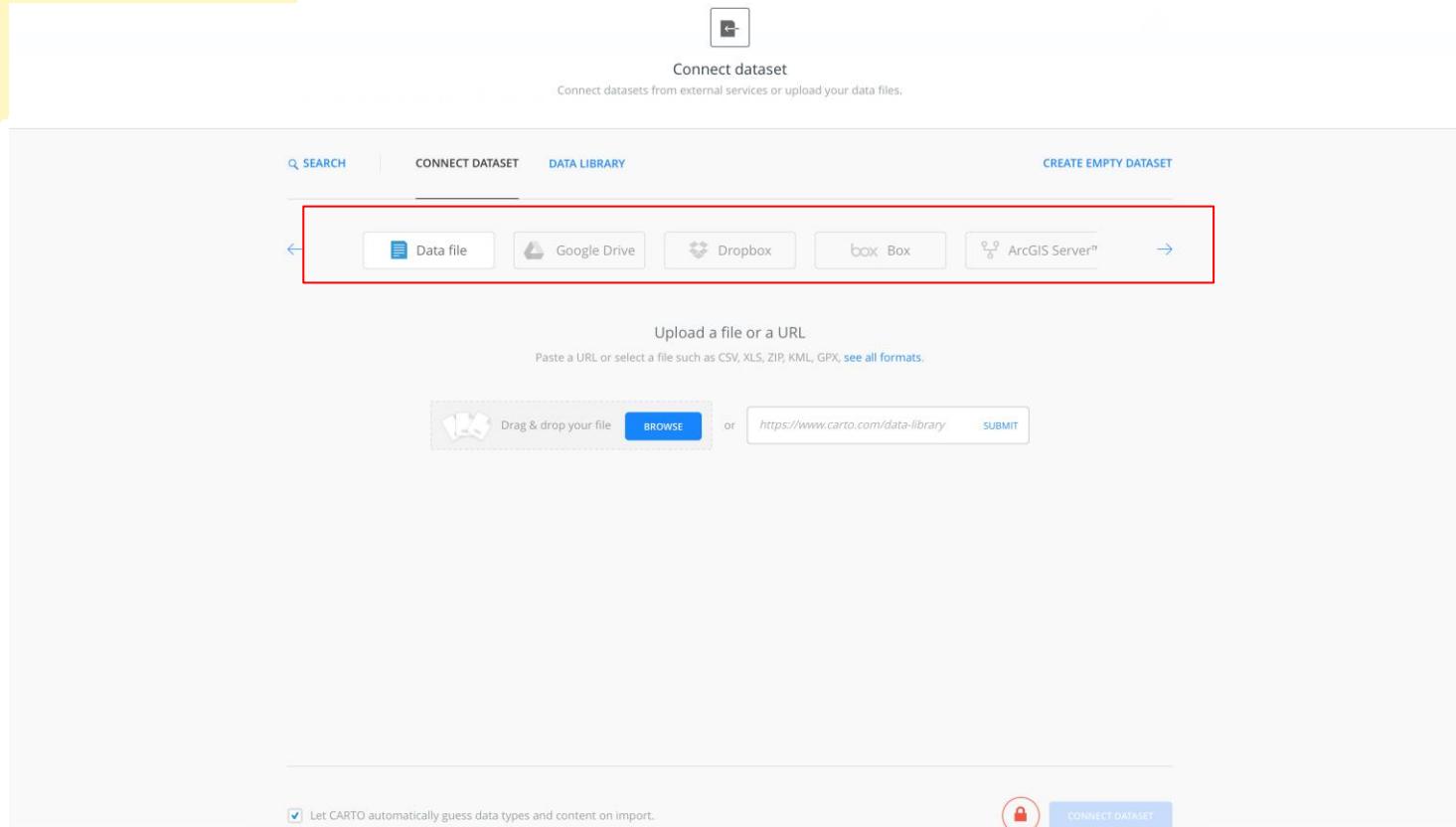
Drag and drop your data file

 Portland Portland building footprints.	632.6K Rows 217 MB 2 minutes ago from <a href="#">Portland Open Data</a> <a href="#">BUILDING FOOTPRINTS</a>
 US Census Tracts US Census Tracts.	73.9K Rows 645 MB 2 minutes ago from <a href="#">US Census</a> <a href="#">US CENSUS</a>
 Chicago Chicago building footprints.	820.9K Rows 242 MB 2 minutes ago from <a href="#">Chicago Open Data</a> <a href="#">BUILDING FOOTPRINTS</a>
 Storm events 2014 Storm event locations in the U.S. 2014.	53.0K Rows 8 MB 2 minutes ago from <a href="#">NOAA's NCDC</a> <a href="#">PHYSICAL DATASETS</a>
 Public use microdata area	2,379 Rows 13 MB 2 minutes ago from <a href="#">US Census</a>

0 datasets selected

CREATE MAP

# Conexión a datos. Datasets



The screenshot shows the 'Connect dataset' page of the CARTO platform. At the top, there's a header with a search bar, a 'CONNECT DATASET' button, a 'DATA LIBRARY' link, and a 'CREATE EMPTY DATASET' link. Below the header, a section titled 'Upload a file or a URL' allows users to paste a URL or select a file from formats like CSV, XLS, ZIP, KML, GPX. There are also links for 'see all formats', 'SUBMIT', and 'BROWSE'. A red box highlights a row of five data source connection buttons: 'Data file', 'Google Drive', 'Dropbox', 'Box', and 'ArcGIS Server'. At the bottom, there's a checkbox for letting CARTO automatically guess data types and content on import, a lock icon, and a 'CONNECT DATASET' button.

# Conexión a datos. Datasets (II)

- Conexión a BBDD
  - Via curl (bajo solicitud)
  - Deshabilitada por defecto



Microsoft®  
SQL Server®



```
curl -v -H "Content-Type: application/json" -d '{\n    "connector": {\n        "provider": "mysql",\n        "connection": {\n            "server": "myserver.com",\n            "database": "mydatabase",\n            "username": "myusername",\n            "password": "mypassword"\n        },\n        "table": "mytable",\n        "encoding": "iso88591"\n    }\n}' "https://{{username}}.carto.com/api/v1/imports/?api_key={{API_KEY}}"
```

## Conexión a datos. Datasets (III)

- Y conexiones de terceros ...



ArcGIS®

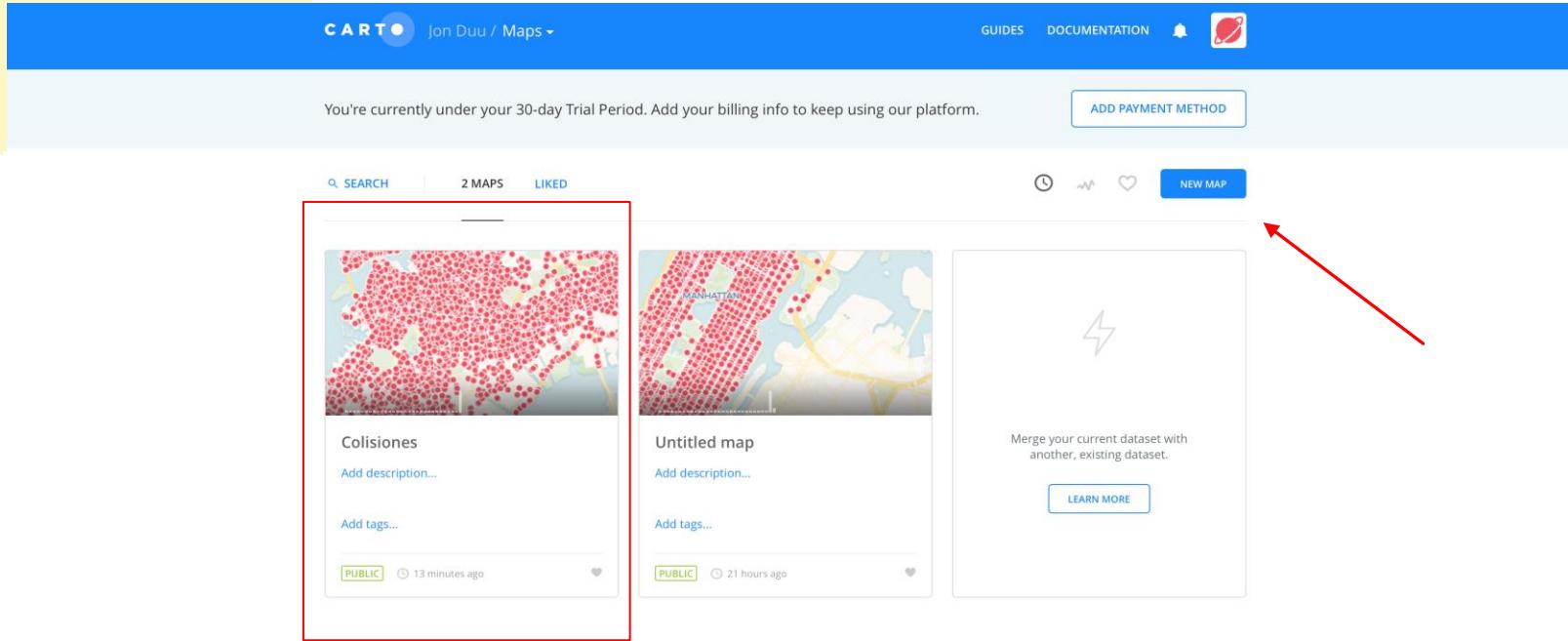
splunk®



Qlik®



# Mapas



The screenshot shows the CARTO platform interface. At the top, there's a blue header bar with the CARTO logo, the user name "Jon Duu / Maps", and navigation links for "GUIDES", "DOCUMENTATION", a notification bell, and a profile icon.

Below the header, a message says: "You're currently under your 30-day Trial Period. Add your billing info to keep using our platform." with a "ADD PAYMENT METHOD" button.

The main area displays three map datasets:

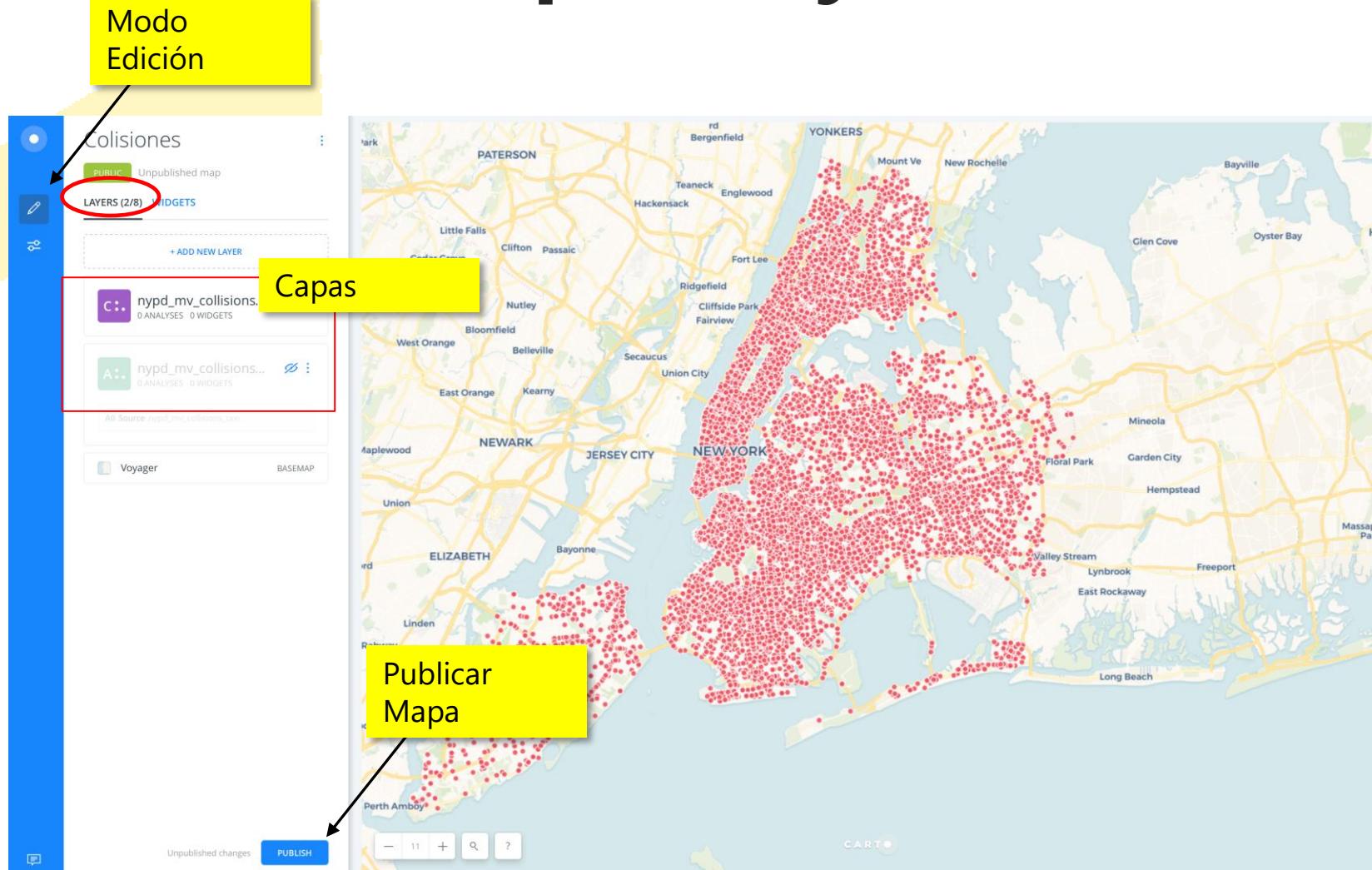
- Colisiones**: A map showing numerous red dots representing collision points across a geographic area. This dataset is highlighted with a red border.
- Untitled map**: A map of Manhattan with red dots overlaid, showing a specific dataset.
- A third map is partially visible on the right.

Below the maps, each dataset has fields for "Add description...", "Add tags...", and status indicators ("PUBLIC", timestamp, and a heart icon). There are also edit and delete icons.

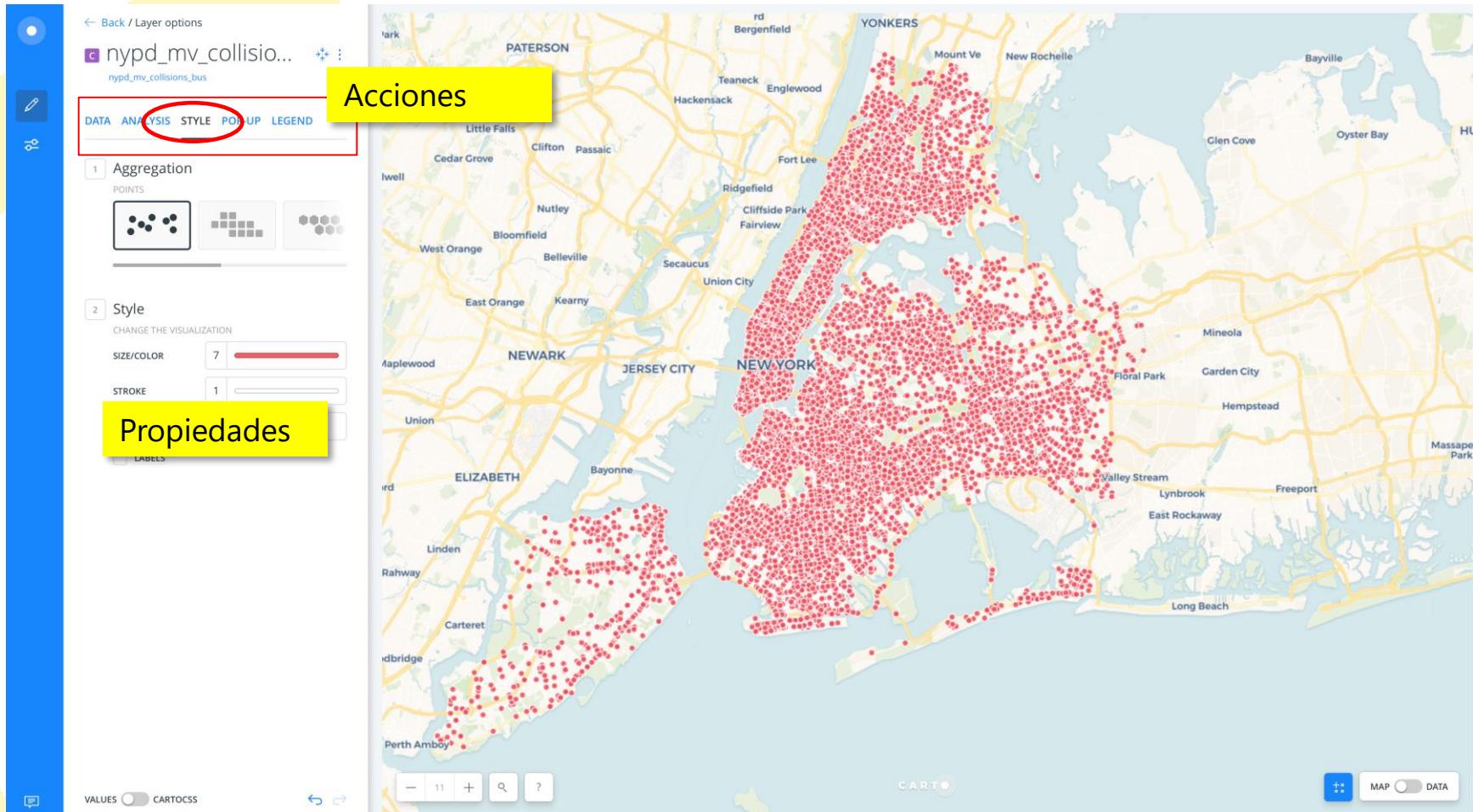
At the top right of the main area, there are icons for a timer, a waveform, a heart, and a "NEW MAP" button. A red arrow points from the "NEW MAP" button towards the third map area.

On the far right, there's a "Merge your current dataset with another, existing dataset." section with a "LEARN MORE" button.

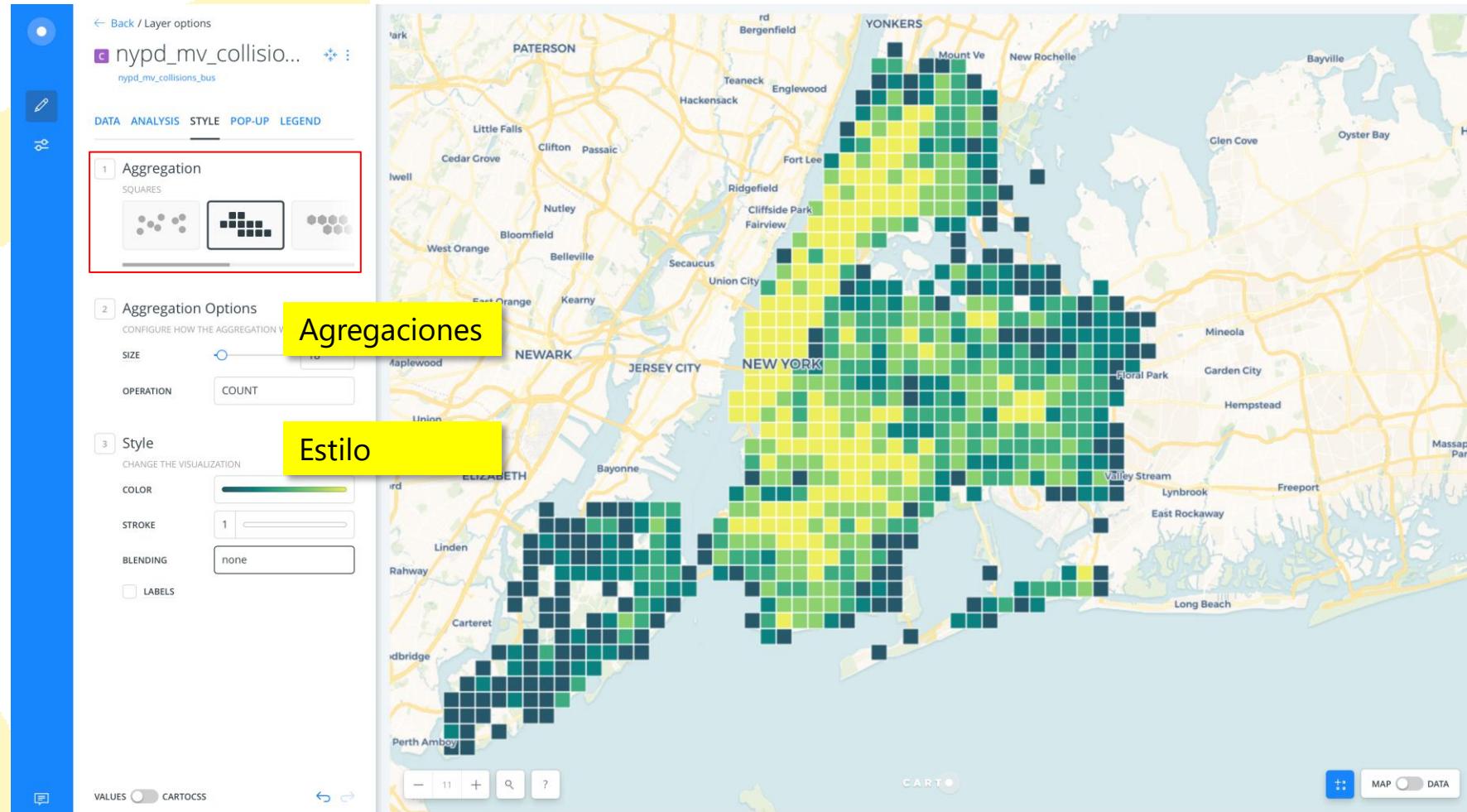
# Mapas. Layers



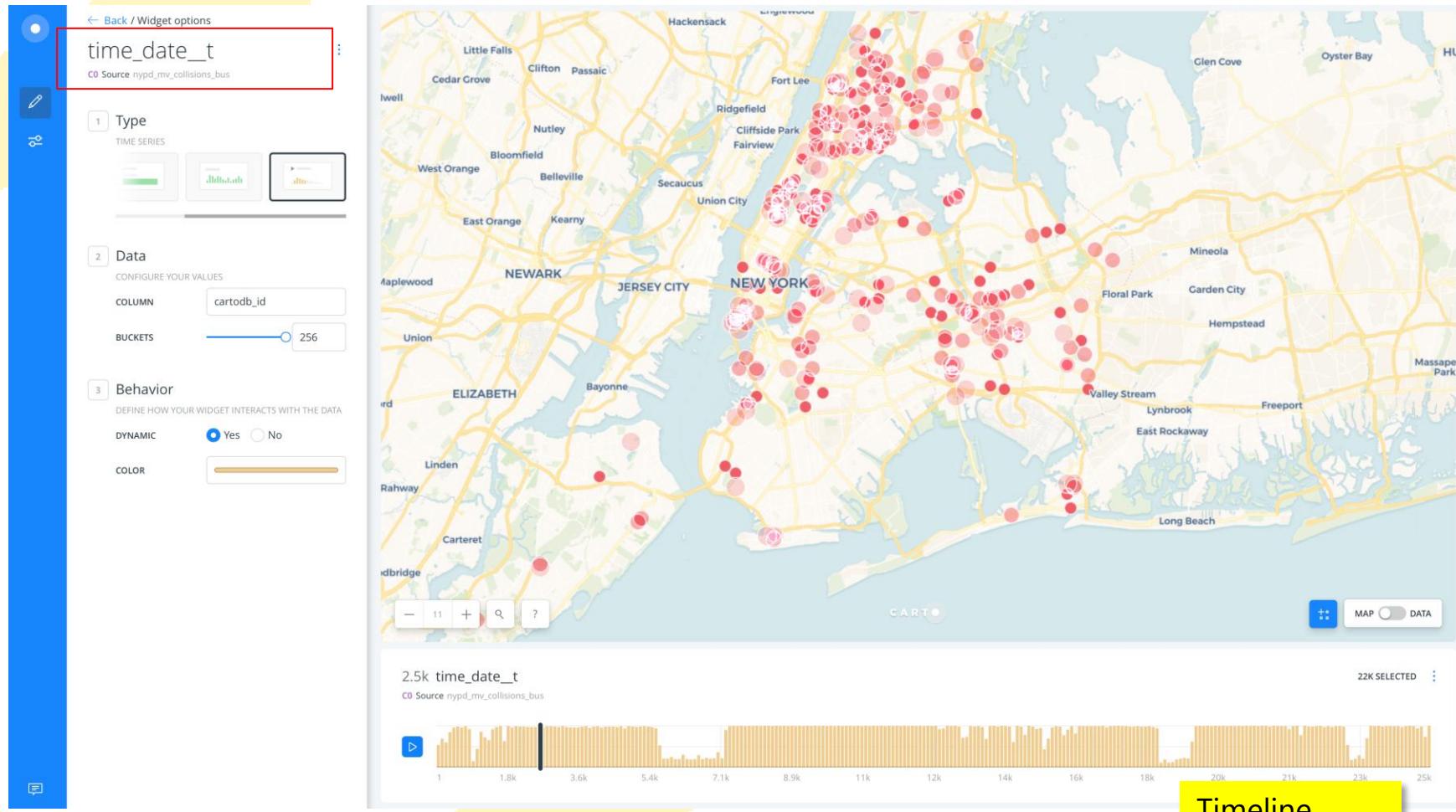
# Layers. Style



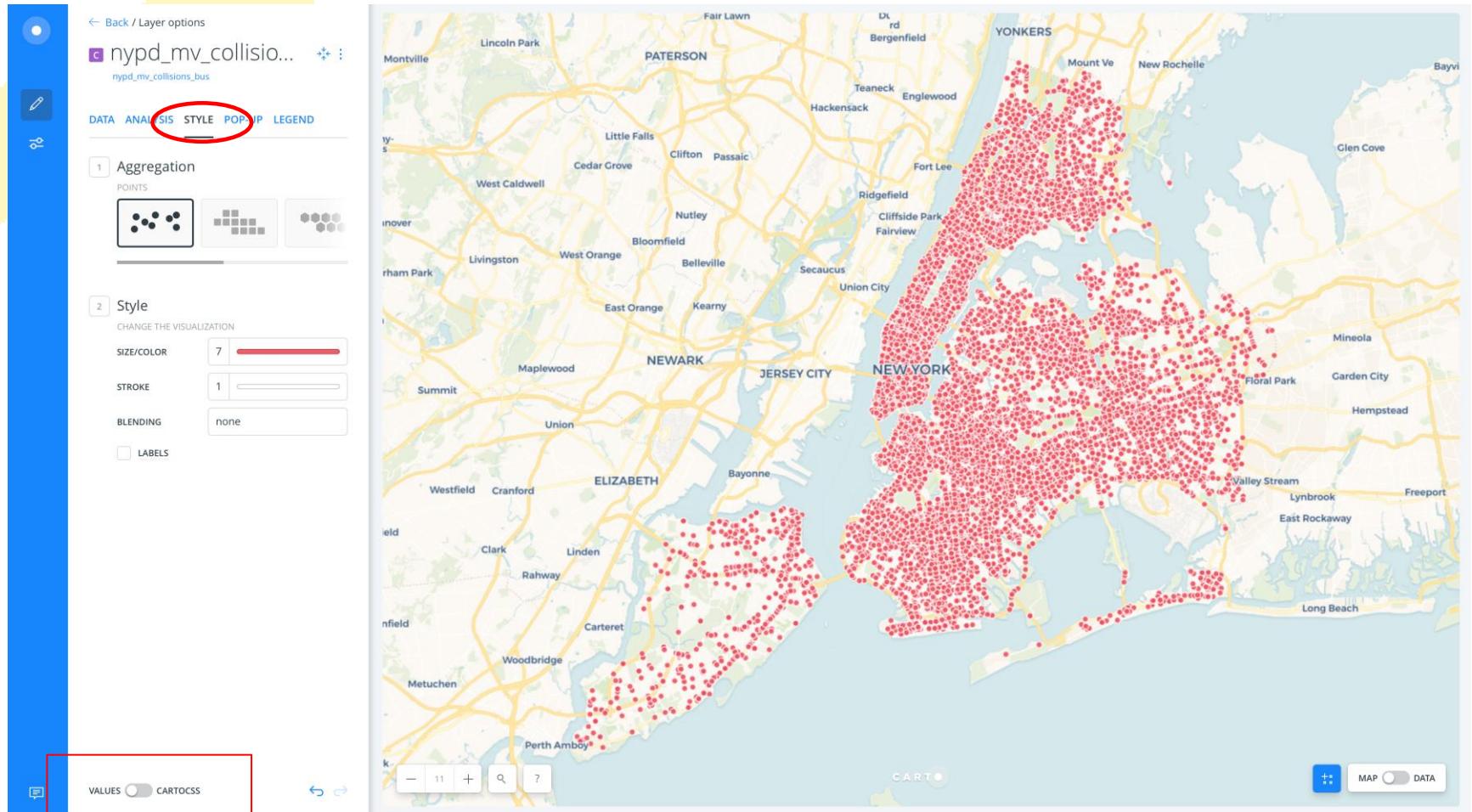
# Layers. Style (II)



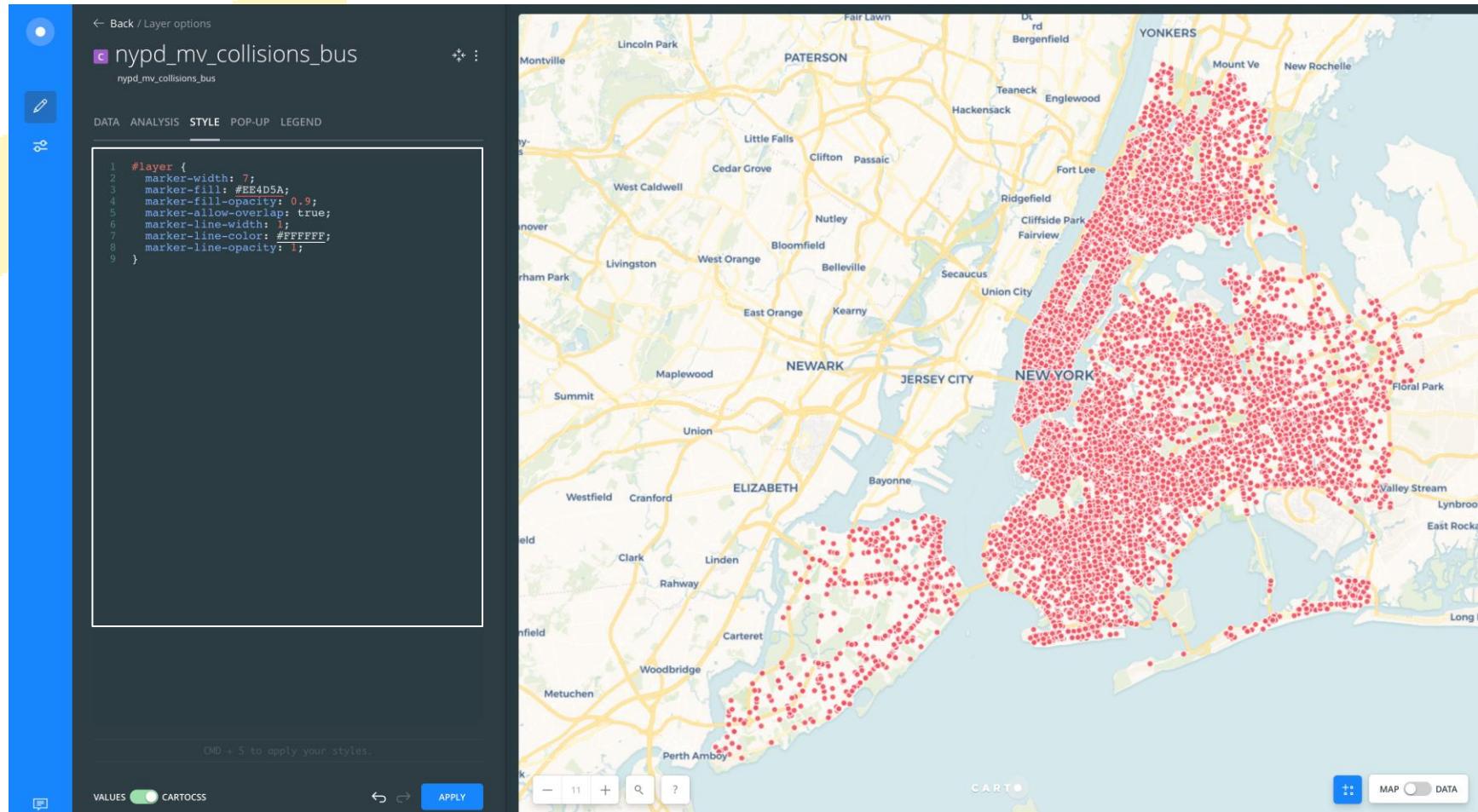
# Layers. Style (III)



# Layers. Style (IV)

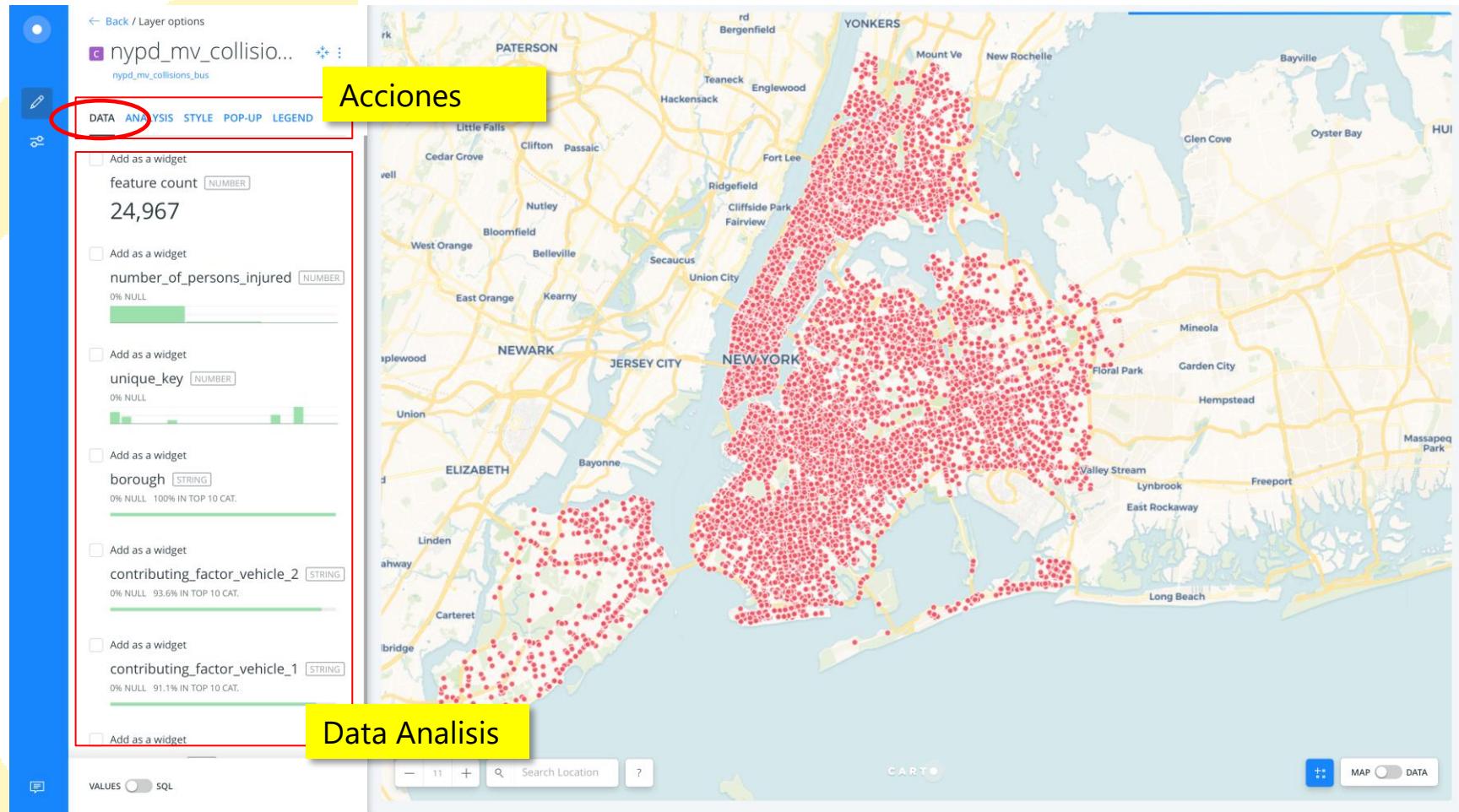


# Layers. Style (V)

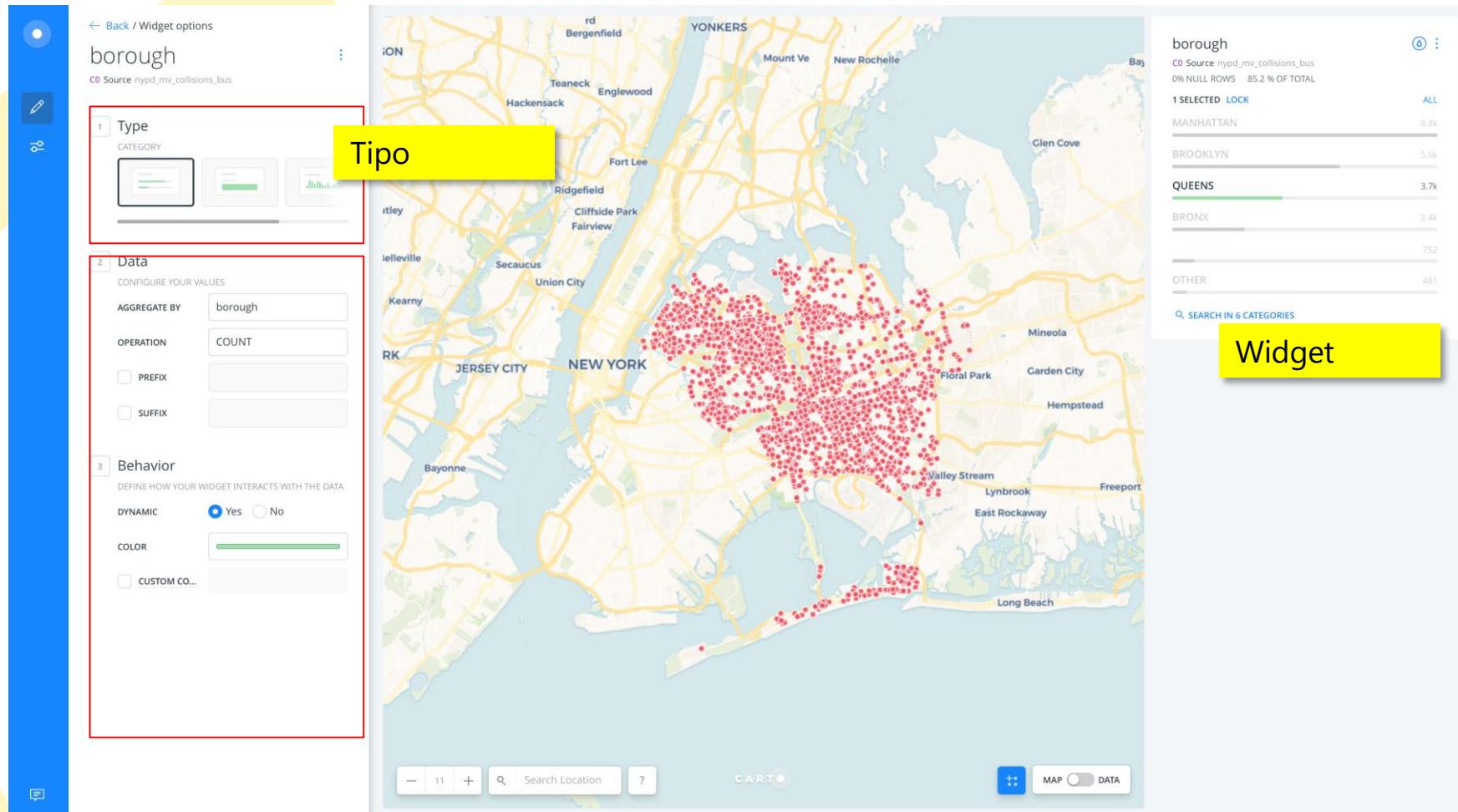


<https://carto.com/docs/cartocss/>

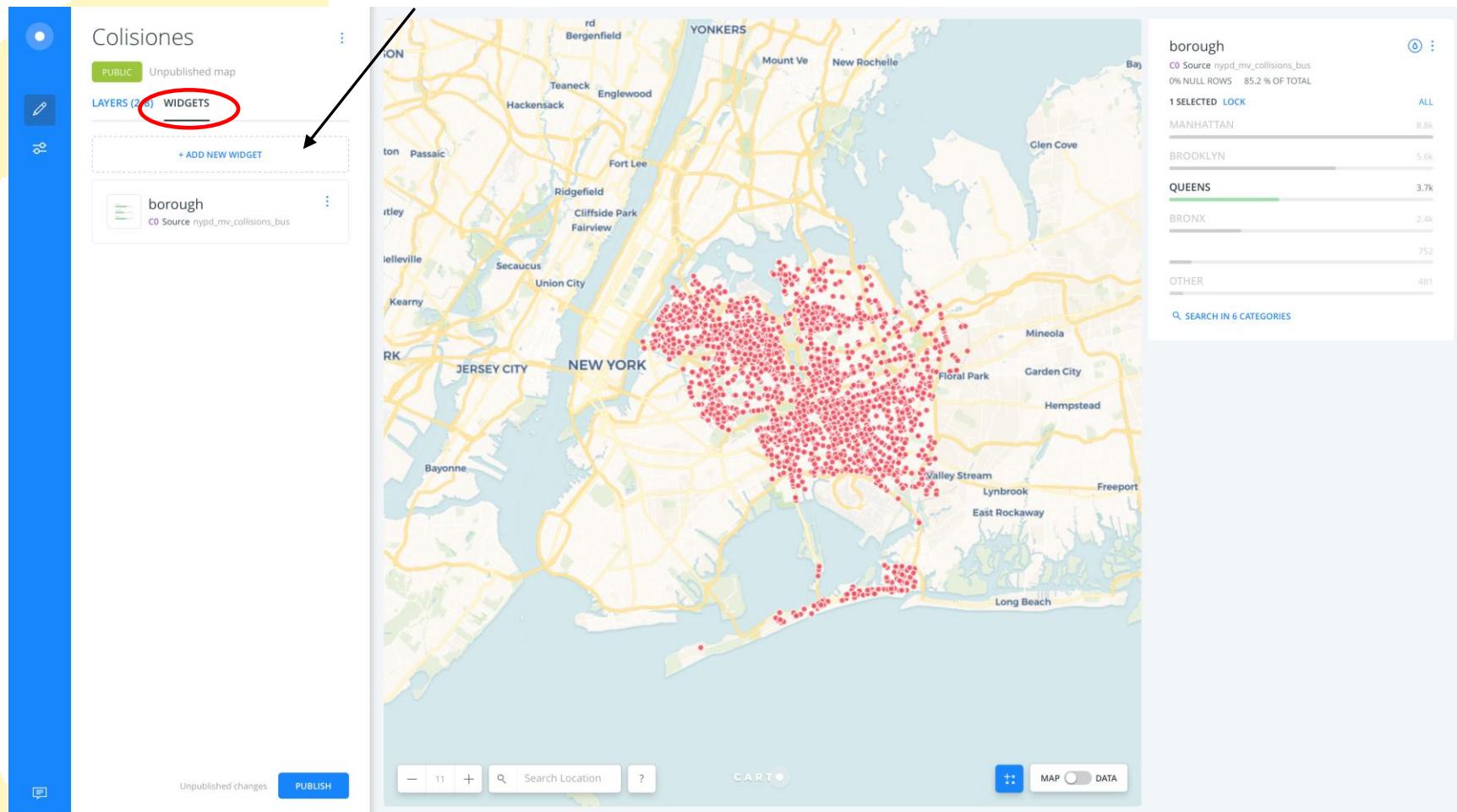
# Layers. Data



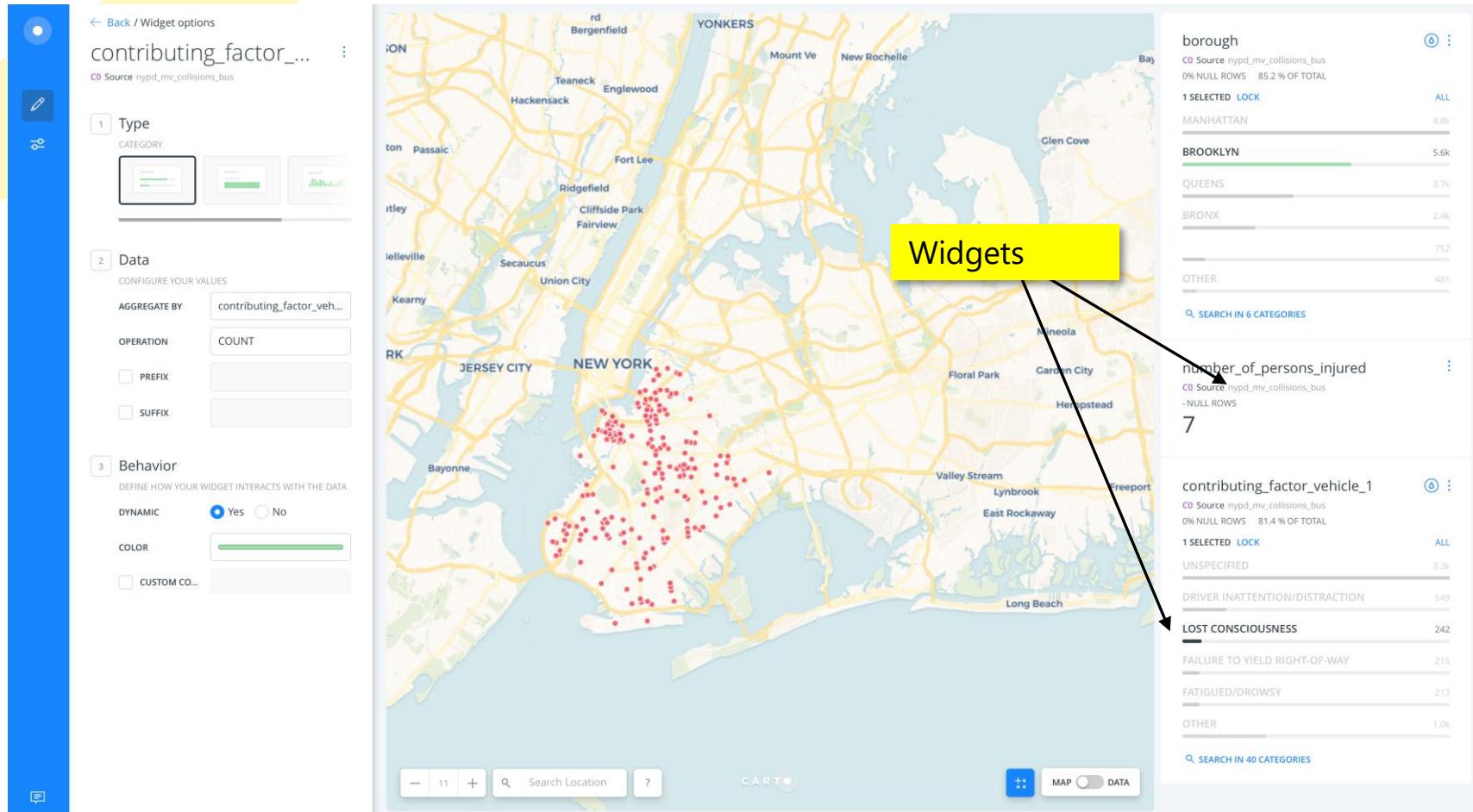
# Layers. Data (II)



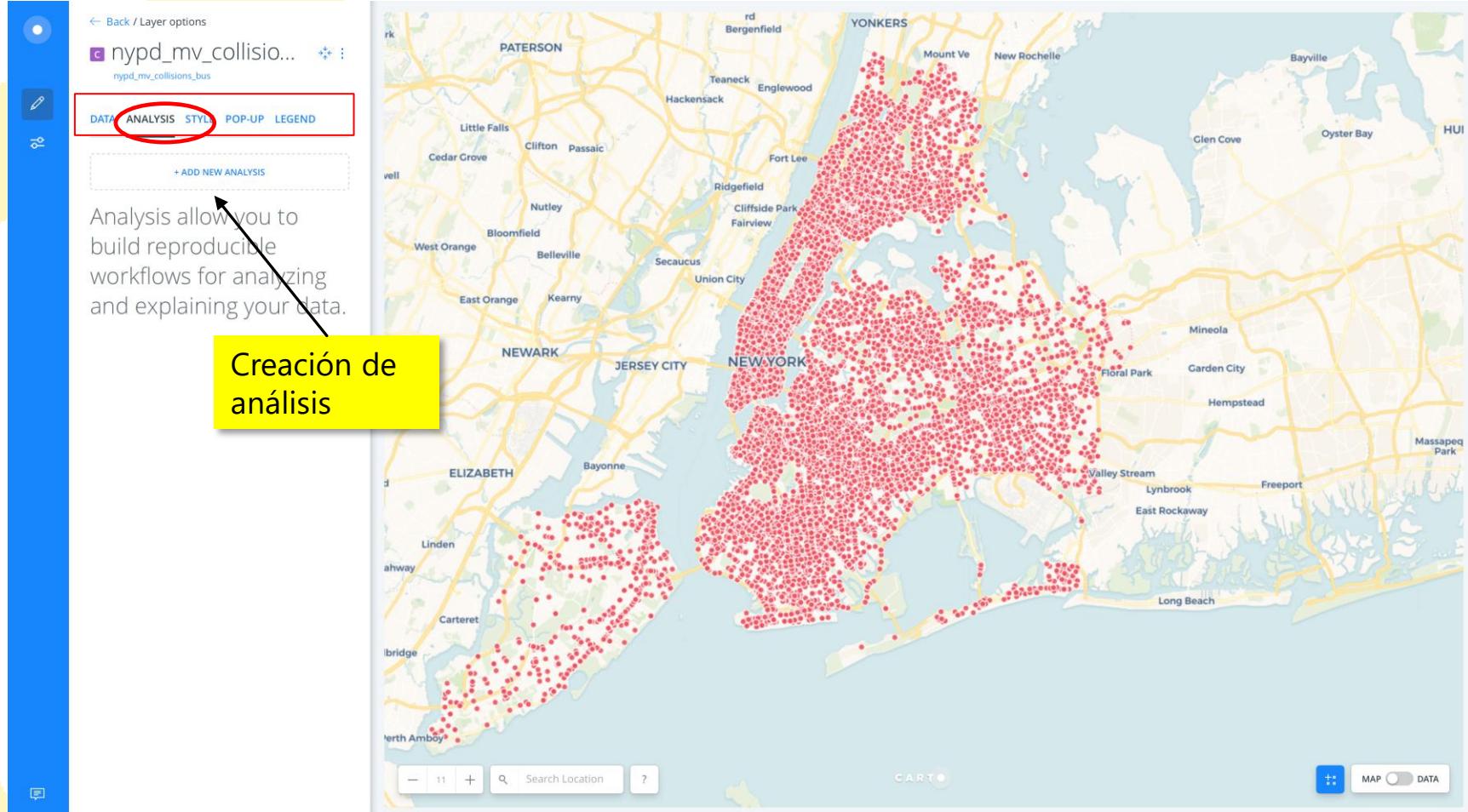
# Layers. Widgets



# Layers. Widgets (II)



# Layers. Analysis



# Layers. Analysis (II)

## Add a new analysis

Select the analysis you want to add

[ALL](#) [CREATE AND CLEAN](#) [ANALYZE AND PREDICT](#) [TRANSFORM](#)



### Intersect second layer

Intersect with a second layer and calculate aggregations on the fly.

[MORE INFO](#)



### Create areas of influence

Use travel time (e.g. walking or driving) or distance to calculate areas of influence from points.

[MORE INFO](#)



### Find centroid of geometries

Calculate a direct or weighted centroid based on all rows of a layer or by categories.

[MORE INFO](#)



### Group points into polygons

Aggregate points into polygons such as convex hulls or bounding boxes.

[MORE INFO](#)



### Filter by column value

Keep or discard rows according to a custom criteria.

[MORE INFO](#)



### Filter points in polygons

Filter points intersecting your polygons layer, and augment those with your polygons columns.

[MORE INFO](#)



### Connect with lines

Create lines from point datasets using a variety of methods.

[MORE INFO](#)



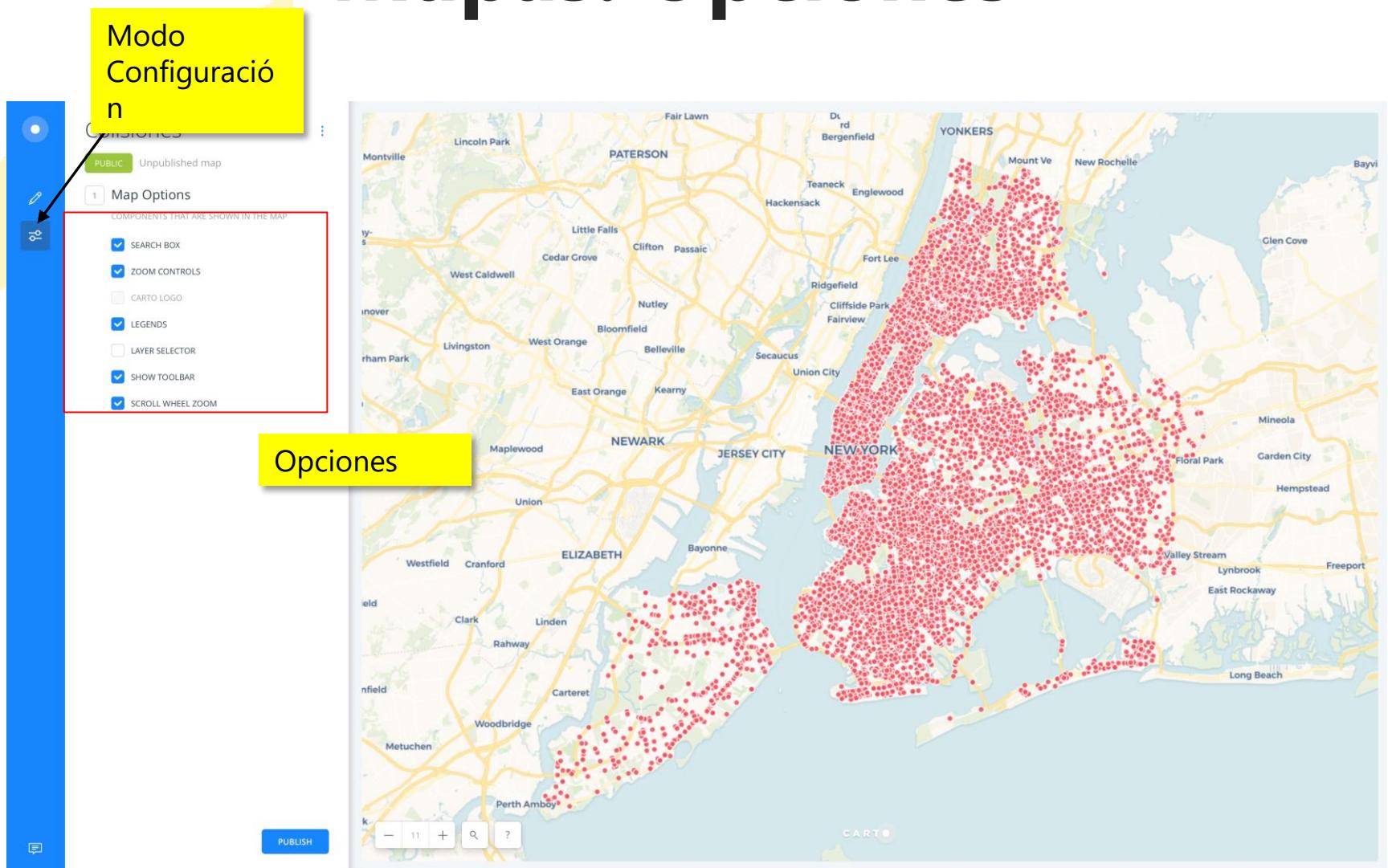
### Subsample percent of rows

Subsample the rows in a dataset based on a specified percent.

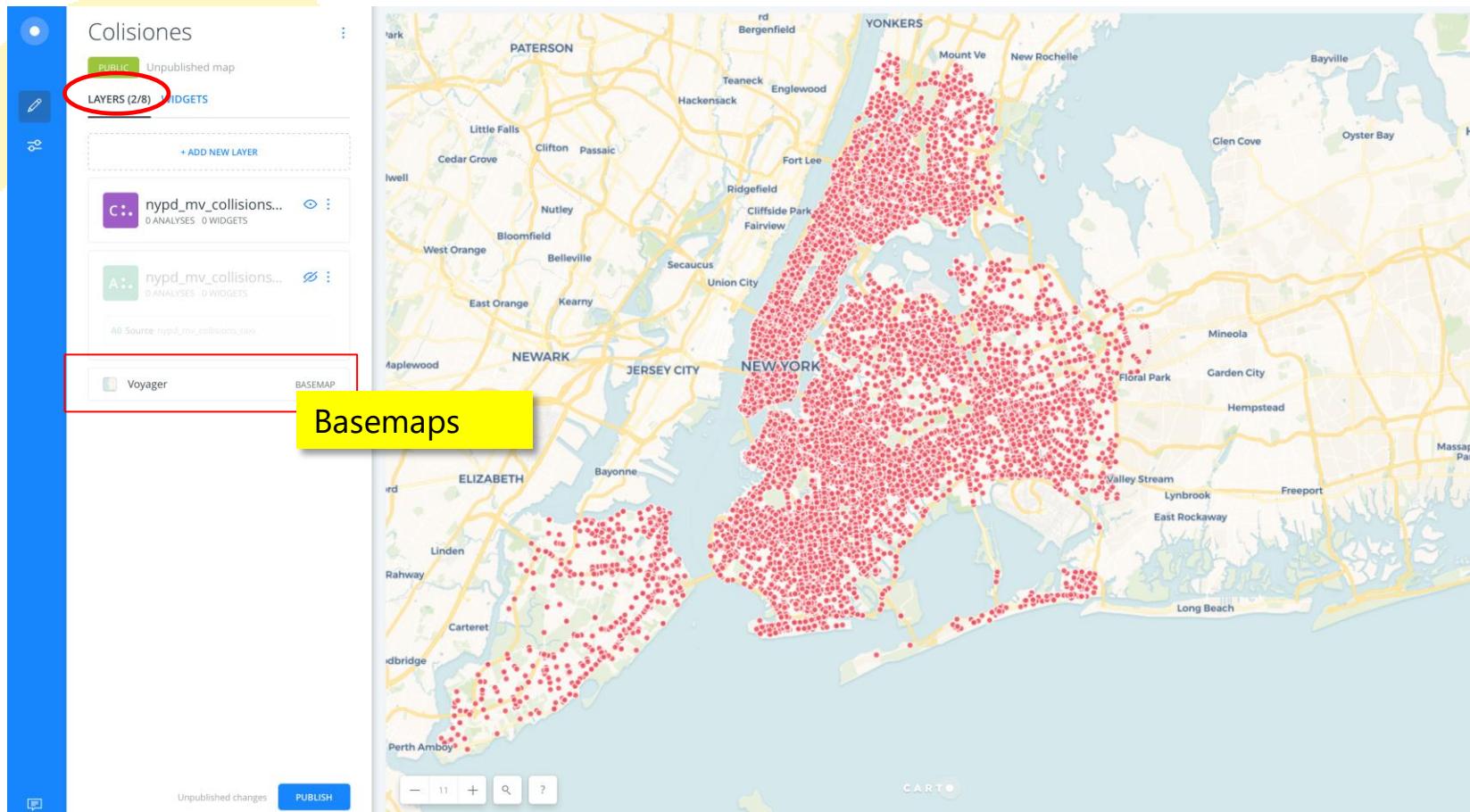
[MORE INFO](#)

[ADD ANALYSIS](#)

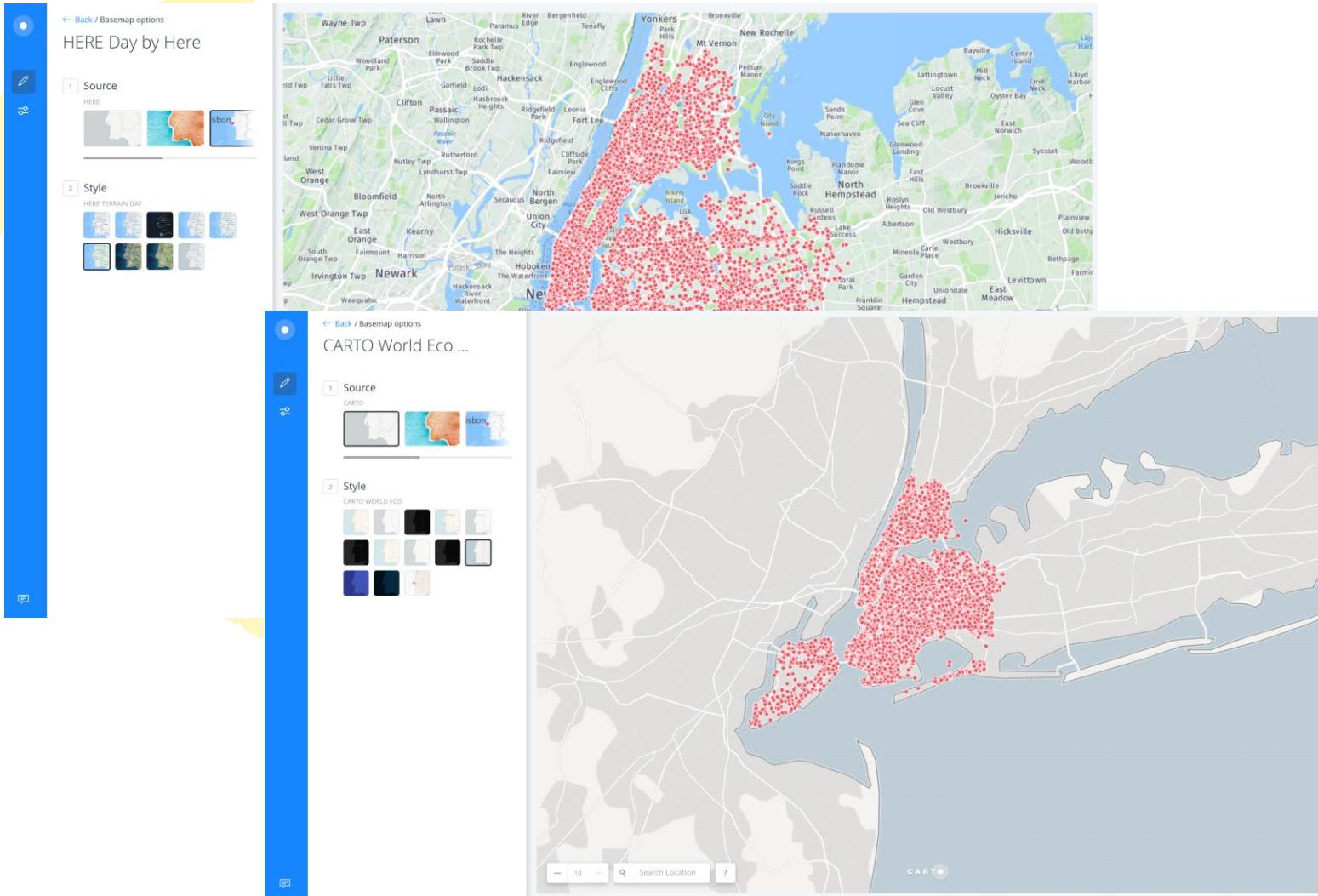
# Mapas. Opciones



# Mapas. Basemaps



## Mapas. Basemaps (II)



# Publicación

← Back

Colisiones

PUBLIC PUBLISH

Never published.



Get the link

Send to your friends, coworkers, or post it in your social networks.



Embed it

Insert your map into your blog, website, or simple application.

[Get a simple URL](#)

Click on Publish to start sharing your map on the web

From the moment you click on publish, you will need to use this window to update your changes on the public version.



Interested in sharing within your organization?

[Contact us](#) to try one of our Enterprise plans

DONE

## Recursos

- Formación
  - <https://carto.com/academy/courses/beginners-course>
  - <https://carto.com/learn/guides>
- Conexión a Twitter
  - <https://carto.com/learn/guides/data-and-sql/connecting-twitter-data>
- Tutorial
  - [https://carto.com/docs/tutorials/electoral\\_map](https://carto.com/docs/tutorials/electoral_map)

# ¿Alguna pregunta?

