

# Data exploration and Visualization

José Antonio Gascón Pinedo  
jagpidp@gmail.com





# ¿ Cómo van a ser estas clases?



- Totalmente participativas y prácticas con Tableau.
- Aunque veremos conceptos con el apoyo de las diapositivas.

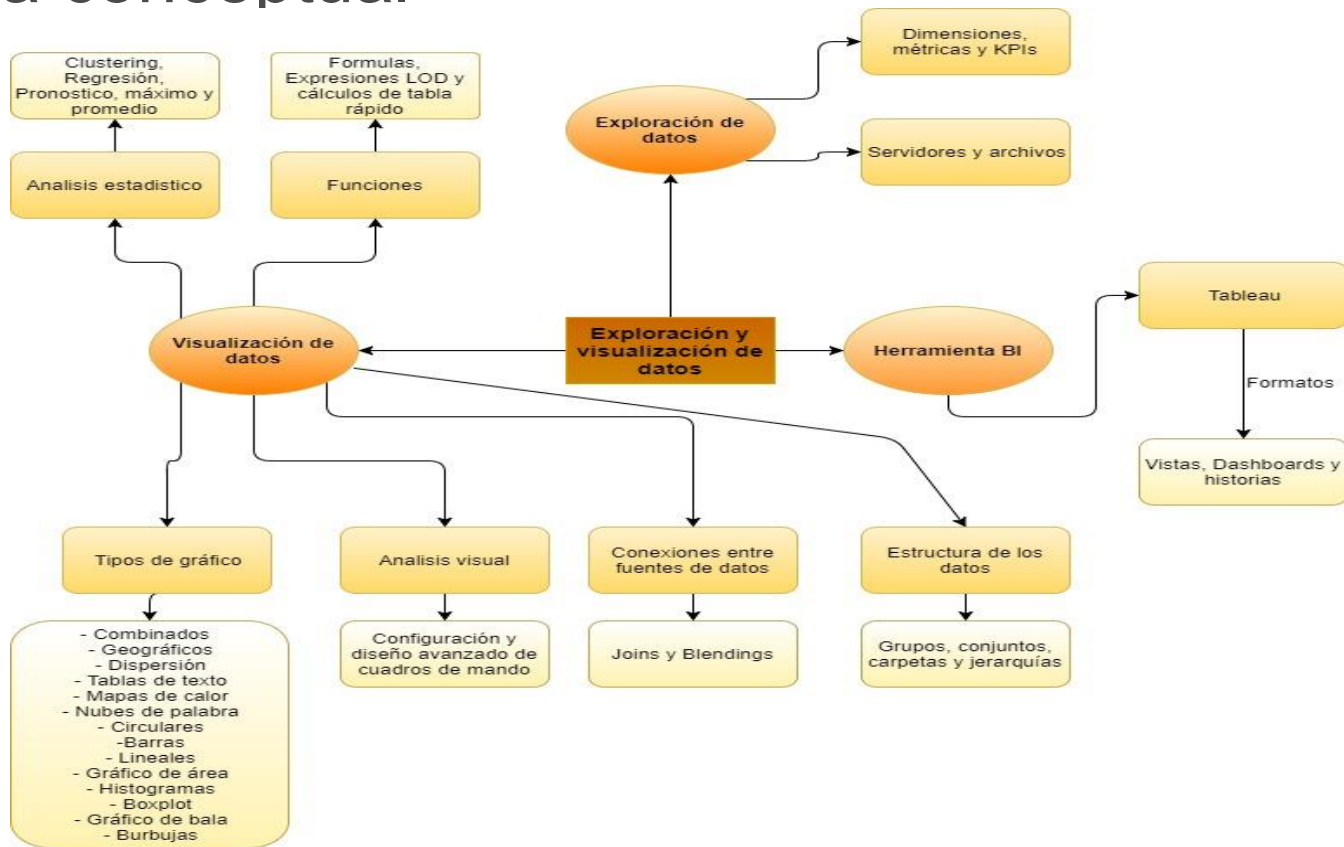
***Consejo: Participar en todos los ejercicios, dar ideas....  
preguntar todo lo que se os ocurra***



# Preparación del entorno de trabajo

- Todos los asistentes deben tener instalado Tableau desktop 2021.1.
- Para los ejercicios prácticos usaremos dos ficheros excel con datos de Airbnb:
  - **airbnb-listing Madrid**
  - **Nvecas alquilado**

# Mapa conceptual

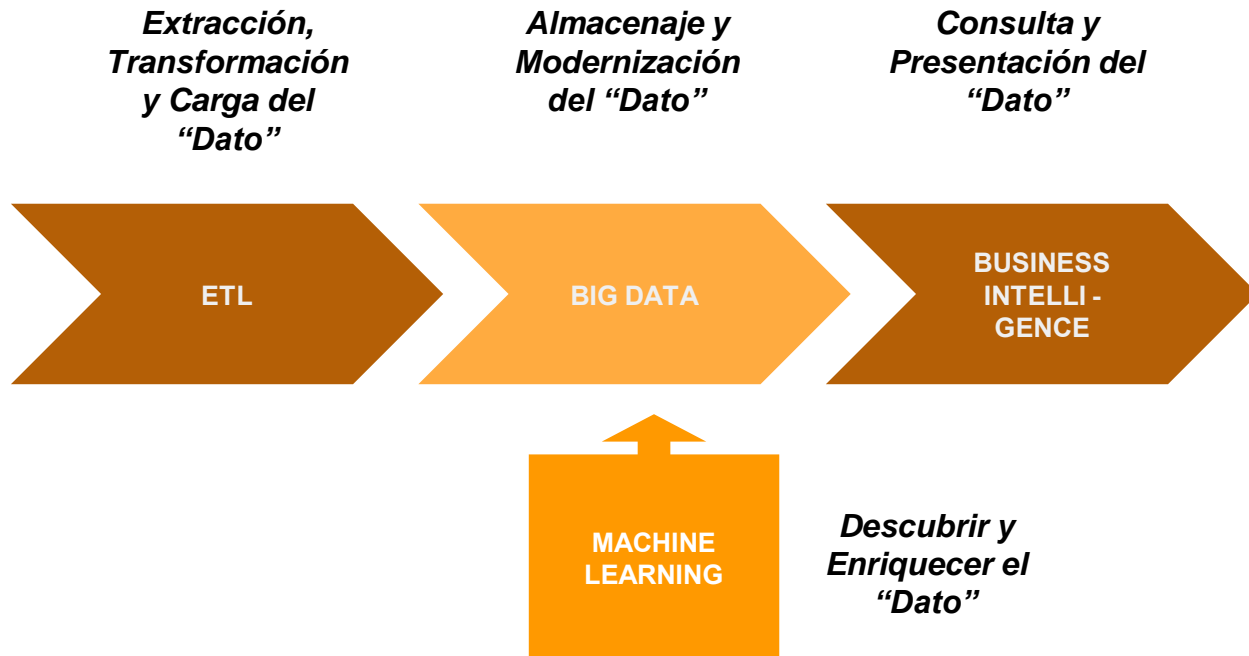


# Índice

1. **Introducción a Tableau**
2. Conexión a datos
3. Simplificación y ordenación de los datos
4. Organización de los datos
5. División de datos por fecha
6. Mapeo geográfico de datos.
7. Panel de análisis
8. Personalización de los datos
9. Cálculos de tablas rápidos
10. Configuración y diseño avanzado de dashboards
11. Visualización de relación entre valores numéricos
12. Visualización de valores específicos
13. Visualización de un desglose de un todo
14. Visualización de distribuciones
15. Análisis visual

# 1. Introducción a Tableau

Arquitectura para el análisis de datos:






# 1. Introducción



# 1. Introducción

## Comparación de proveedores de BI

	MICROSOFT POWER BI	QLIK SENSE	TABLEAU
<b>Capacidades visuales</b> 	<ul style="list-style-type: none"><li>■ La herramienta más simple de usar entre los proveedores principales de BI, de acuerdo con el "Cuadrante Mágico para Inteligencia de Negocios y Plataformas de Analítica" 2017, de Gartner.</li><li>■ Integra y visualiza datos de fuentes dispares, ya sean basadas en nube, locales o Hadoop.</li></ul>	<ul style="list-style-type: none"><li>■ Motor de datos en memoria deja a los usuarios construir gráficos interactivos. Guardar datos en memoria permite actualizar gráficos rápidamente mientras los usuarios exploran visualmente los conjuntos de datos.</li><li>■ Motor de datos robusto permite combinar y visualizar múltiples fuentes de datos.</li></ul>	<ul style="list-style-type: none"><li>■ Herramienta de visualización más atractiva e intuitiva, según el "Cuadrante Mágico de BI y plataformas analíticas 2017 de Gartner.</li><li>■ Procesamiento de datos en memoria para refrescos rápidos en visualizaciones al explorar datos. Conjuntos de datos más grandes requieren llamadas directas a fuente.</li></ul>
<b>Capacidades avanzada de analítica</b> 	<ul style="list-style-type: none"><li>■ Soporta visualizaciones basadas en lenguaje R, incluyendo previsión, agrupamiento y árboles de decisión.</li></ul>	<ul style="list-style-type: none"><li>■ No hay soporte para R o Python, pero la compañía dice que esas características ya vienen.</li><li>■ Analítica predictiva, clustering y regresiones son solo posibles a través de conexiones API con software de terceros.</li></ul>	<ul style="list-style-type: none"><li>■ Soporte totalmente integrado para lenguajes R y Python.</li><li>■ Herramientas nativas para clustering y previsiones.</li></ul>
<b>Disponibilidad de nube</b> 	<ul style="list-style-type: none"><li>■ Software primero nube ofrecido a través de la plataforma Azure de Microsoft.</li><li>■ Opción de escritorio disponible. Cuentas de nube son necesarias para compartir visualizaciones.</li></ul>	<ul style="list-style-type: none"><li>■ Ofrece producto de nube SaaS completamente gestionado. La mayoría de los clientes escoge ejecutar la versión de servidor.</li></ul>	<ul style="list-style-type: none"><li>■ Puede desplegarse en la nube gestionada por Tableau o en plataformas terceras, incluyendo Amazon Web Services y Microsoft Azure.</li></ul>
<b>Precio</b> 	<ul style="list-style-type: none"><li>■ \$9.99 por usuario, por mes.</li></ul>	<ul style="list-style-type: none"><li>■ Enterprise: \$1,500 por token; 1 token compra uso ilimitado para un usuario o 10 pases temporales de inicio de sesión.</li><li>■ Nube: \$20 al mes por un usuario, \$25 al mes, por usuario para múltiples inicios de sesión.</li></ul>	<ul style="list-style-type: none"><li>■ Tableau Desktop Personal: \$35 por usuario, al mes.</li><li>■ Tableau Desktop Professional: \$70 por usuario, al mes.</li><li>■ Tableau Server: \$35 por usuario, al mes.</li><li>■ Tableau Online: \$42 por usuario, al mes.</li></ul>
<b>Letra pequeña</b> 	<ul style="list-style-type: none"><li>■ Límite de suscripción: 10 GB almacenamiento nube para datos.</li><li>■ Puede requerir costos adicionales para escalar la capacidad de datos.</li></ul>	<ul style="list-style-type: none"><li>■ Límite de suscripción de Qlik Sense Cloud Business: 500 GB de almacenamiento de nube por grupo de trabajo.</li></ul>	<ul style="list-style-type: none"><li>■ Límite de suscripción de Tableau Online: 100 GB de almacenamiento de datos en la nube.</li></ul>

SEOF TECHNOLOGIES ALL RIGHTS RESERVED TechTarget



# 1. Introducción a Tableau

Tableau ayuda a las personas a  
ver y comprender sus datos

# 1. Introducción a Tableau

Creemos en la idea de  
otorgar  
el poder a las personas

Para crear una empresa que cambie completamente la forma en que las personas ven y comprenden los datos, hay que aplicar una filosofía diferente. Los fundadores de Tableau empaparon a su compañía con puntos de vista revolucionarios.

Liberar los datos

Capacitar a las personas

Diseñar para las personas

# 1. Introducción a Tableau

## Liberar los datos

Creemos que el análisis de datos debería consistir en hacer preguntas, no en aprender a usar un software. Tableau crea productos inspiradores y fáciles de usar que ayudan a las personas a crecer con sus datos. De este modo, los datos puedan alcanzar su mayor potencial y tener un impacto positivo en el mundo.

## Capacitar a las personas

Cuando una empresa ofrece herramientas de análisis de autoservicio a sus empleados, estos se sienten distintos. Respetados. Capaces. Poderosos. Comienzan a impulsar a la organización de maneras inimaginables. Expresan todo su ingenio y creatividad.

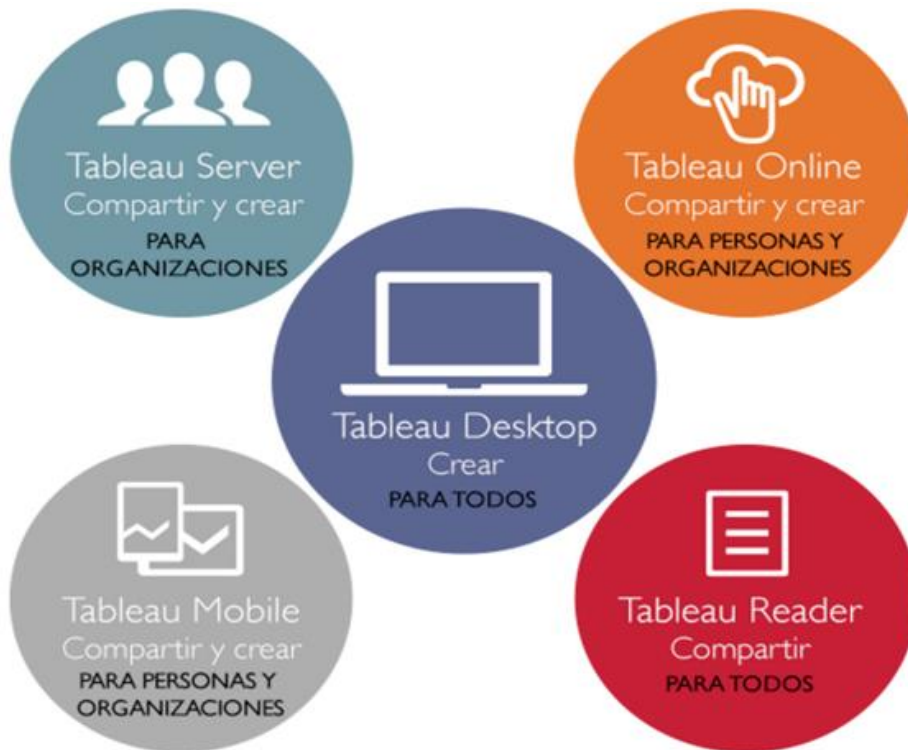
## Diseñar para las personas

Diseñamos nuestros productos concentrándonos en las personas. En aquellas que trabajan en grandes o pequeñas empresas, en las que brindan servicios al gobierno o incluso en las que dedican sus esfuerzos a organizaciones sin fines de lucro. Blogueros. Estudiantes. Cualquier persona. Todos.



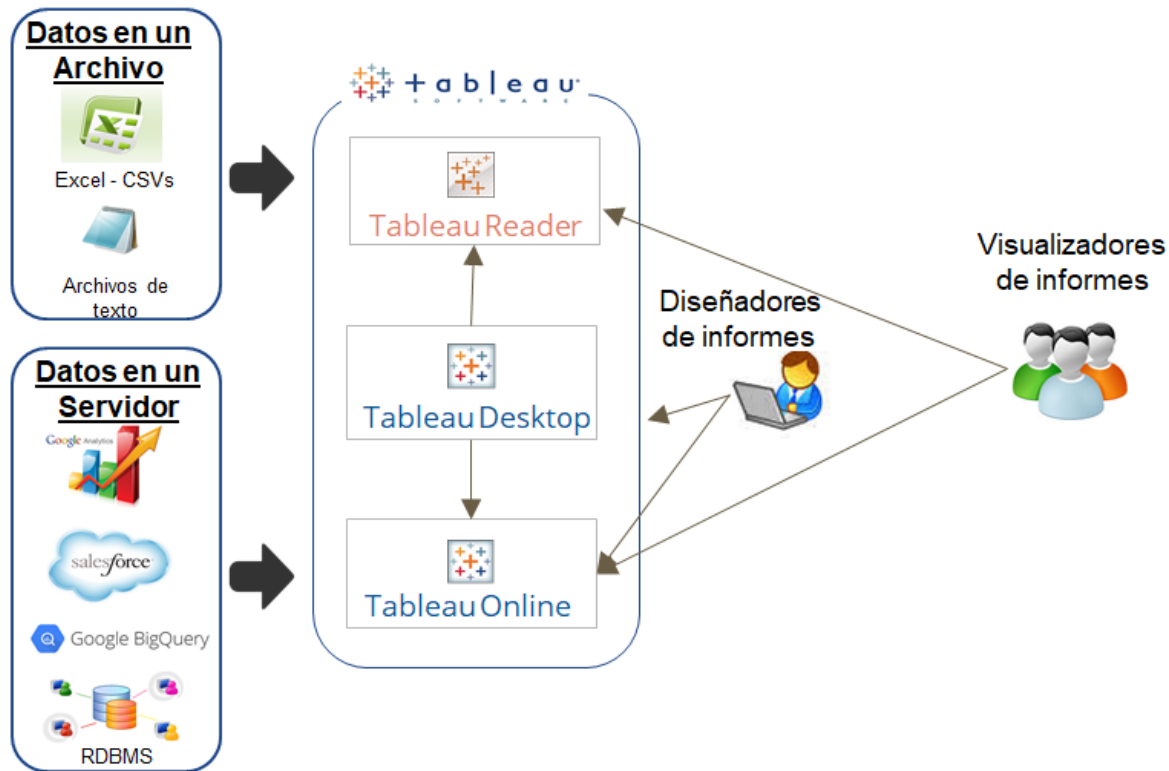
# 1. Introducción a Tableau

## Componentes de Tableau Software



# 1. Introducción a Tableau

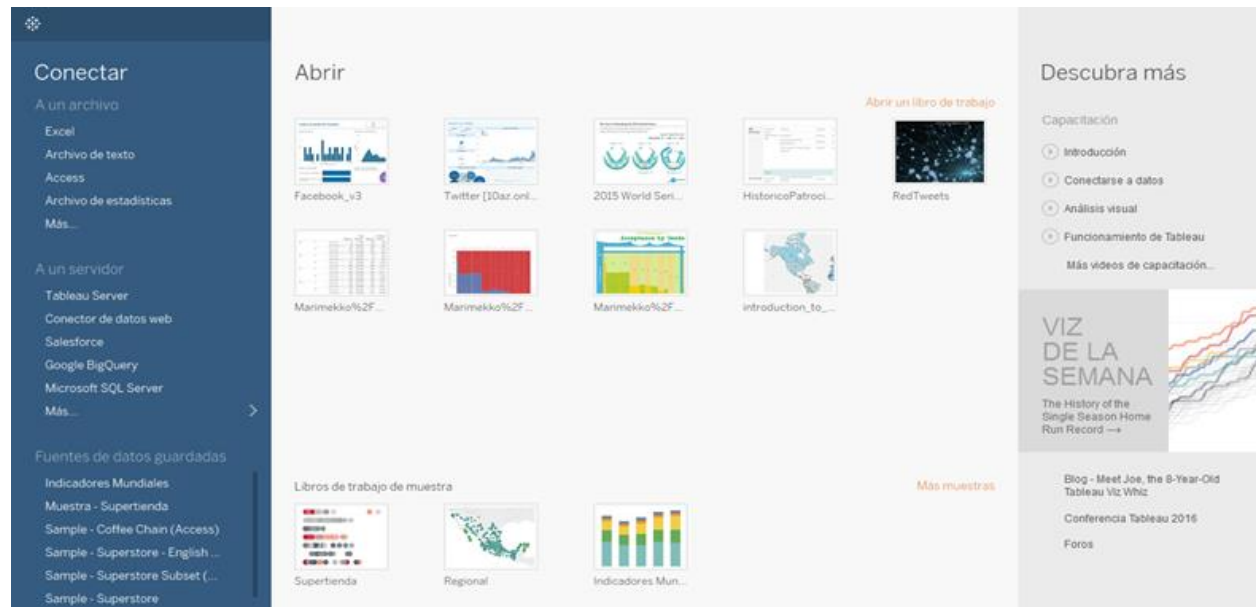
## Arquitectura:



# 1. Introducción a Tableau

## Ventana de inicio de Tableau Desktop

- Creación de conexiones
- Conectar a fuentes guardadas
- Abrir libros de trabajo guardados



# Índice

1. Introducción a Tableau
2. **Conexión a datos**
3. Simplificación y ordenación de los datos
4. Organización de los datos
5. División de datos por fecha
6. Mapeo geográfico de datos.
7. Panel de análisis
8. Personalización de los datos
9. Cálculos de tablas rápidas
10. Configuración y diseño avanzado de dashboards
11. Visualización de relación entre valores numéricos
12. Visualización de valores específicos
13. Visualización de un desglose de un todo
14. Visualización de distribuciones
15. Análisis visual

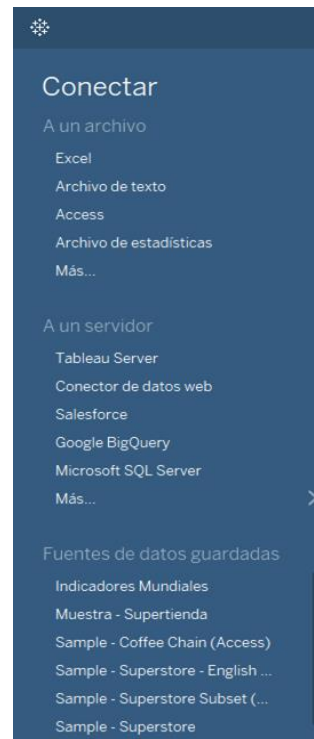


## 2. Conexión a datos

### Pasos para la conexión a datos

En la ventana de inicio de Tableau Desktop, seleccionar un origen de datos entre todos los disponibles:

- Ficheros: Excel, Access, ficheros de texto
- Bases de datos relacionales: SQL Server, Oracle, PostgreSQL...
- Bases de datos en la nube: Google BigQuery, Google Analytics, Salesforce...
- Según el origen de datos:
  - Fichero: seleccionar un fichero en el equipo local
  - Base de datos: proporcionar host, usuario y contraseña





## 2. Conexión a datos

Conexión en vivo  
o con extractos

Tablas de  
los  
informes

Tipo de unión y  
campo de relación

Previsualización  
de los datos

Orders+ (Sample - Superstore)

Conectado a Excel

Libro de trabajo

Sample - Superstore.xlsx

Hojas

Escriba el nombre de la hoja

Orders

People

Returns

Conexión

En vivo

Extracto

Filtros 0 Agregar...

Mostrar alias

Mostrar campos ocultos

Filas 9,994

Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	Country	City	State	Postal Code	Region	Product ID
1	CA-2013-152156	09/11/2013	12/11/2013	Second Class	CG-12520	Claire Gite	Consumer	United States	Henderson	Kentucky	42420	South	FUR-BO-10001798
2	CA-2013-152156	09/11/2013	12/11/2013	Second Class	CG-12520	Claire Gite	Consumer	United States	Henderson	Kentucky	42420	South	FUR-CH-10000454
3	CA-2013-138688	13/06/2013	17/06/2013	Second Class	DV-13045	Darin Van Huff	Corporate	United States	Los Angeles	California	90036	West	OFF-LA-10000240
4	US-2012-108966	11/10/2012	18/10/2012	Standard Class	SO-20335	Sean O'Donnell	Consumer	United States	Fort Lauderdale	Florida	33311	South	FUR-TA-10000577
5	US-2012-108966	11/10/2012	18/10/2012	Standard Class	SO-20335	Sean O'Donnell	Consumer	United States	Fort Lauderdale	Florida	33311	South	OFF-ST-10000760
6	CA-2011-115812	09/06/2011	14/06/2011	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	FUR-FU-10001487
7	CA-2011-115812	09/06/2011	14/06/2011	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	OFF-AR-10002833
8	CA-2011-115812	09/06/2011	14/06/2011	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	TEC-PH-10002275
9	CA-2011-115812	09/06/2011	14/06/2011	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	OFF-BI-10003910
10	CA-2011-115812	09/06/2011	14/06/2011	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	OFF-AP-10002892
11	CA-2011-115812	09/06/2011	14/06/2011	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	FUR-TA-10001539
12	CA-2011-115812	09/06/2011	14/06/2011	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	TEC-PH-10002033
13	CA-2014-114412	16/04/2014	21/04/2014	Standard Class	AA-10480	Andrew Allen	Consumer	United States	Concord	North Caroli...	28027	South	OFF-PA-10002385

Fuente de datos

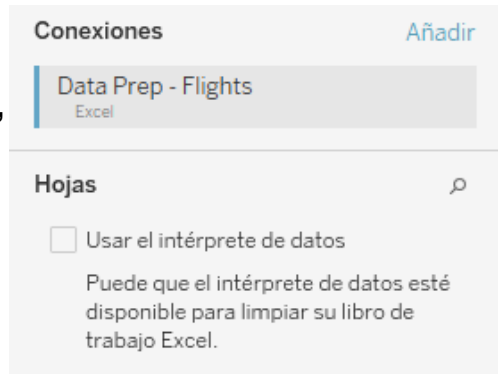
Hoja 1

- Nota: Trabajar con extractos mejora el rendimiento y permite trabajar sin conexión.

## 2. Conexión a datos

### Intérprete de datos

- Permite formatear de manera automática hojas de Excel con estructuras que Tableau no lee correctamente (sub-tablas, encabezados anidados, filas o columnas en blanco, etc.)
- Se activa seleccionando la casilla “Usar el intérprete de datos”
- El intérprete elimina la información extraña de la fuente de datos y actualiza automáticamente la cuadrícula de datos con su interpretación.
- Es posible revisar las transformaciones hechas por Tableau haciendo clic en el vínculo “Revisar los resultados”



# Ejercicios



Abrir Tableau y crear un nuevo libro.  
Conectarse a la hoja excel y observar los campos cargados.

# Índice

1. Introducción a Tableau
2. Conexión a datos
3. **Simplificación y ordenación de los datos**
4. Organización de los datos
5. División de datos por fecha
6. Mapeo geográfico de datos.
7. Resaltar datos con líneas de referencia
8. Personalización de los datos
9. Cálculos de tablas rápidos
10. Configuración y diseño avanzado de dashboards
11. Visualización de relación entre valores numéricos
12. Visualización de valores específicos
13. Visualización de un desglose de un todo
14. Visualización de distribuciones
15. Análisis visual

# 3. Simplificación y ordenación de datos

## Entorno de trabajo de Tableau Desktop



### 3. Simplificación y ordenación de datos

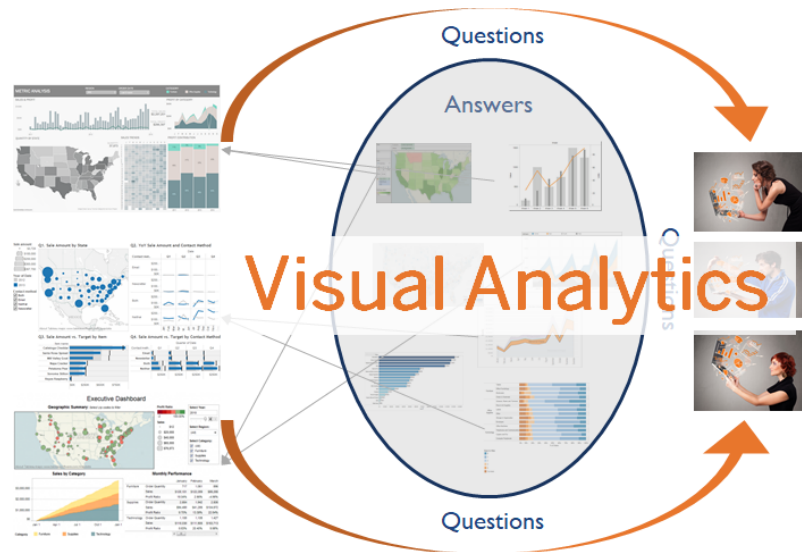
## Dimensiones, Métricas y KPIs

- Dimensiones: atributos, campos no numéricos. Por ejemplo, IDs o fechas.
- Medidas o métricas: campos con valores numéricos. Por ejemplo, compras o ventas
- KPIs: Son las métricas/cálculos necesarios para llegar al objetivo que se ha pensado tras analizar la información.

# 3. Simplificación y ordenación de datos

## ¿Cómo trabajar con mis datos?

1. Definición de objetivos deseados
2. Selección de KPIs adecuadas
3. Medición de resultados
4. Análisis de datos
5. Implementación de mejoras



# 3. Simplificación y ordenación de datos

## Tipos de datos

Icono	Descripción
Abc	Valores de texto
📅	Valores de fecha
🕒	Valores de fecha y hora
#	Valores numéricos
T F	Valores booleanos
🌐	Valores geográficos



# Ejercicios



Entender las dimensiones y las medidas que se han creado en Tableau.

Es importante saber cómo tenemos los datos para poder hacer informes y análisis.

# 3. Simplificación y ordenación de datos

## Creación de vistas

- Para crear una vista se arrastran los campos a los contenedores de filas o columnas.
- Para añadir información a la vista creada se arrastran campos a los contenedores de Páginas, Filtros o Marcas.
- Dentro de la pestaña de Marcas se puede escoger:
  - Tipo de gráfico que se usará
  - Características para ese gráfico:
    - Color
    - Forma
    - Etiqueta
    - Detalle
    - Descripción emergente

### 3. Simplificación y ordenación de datos



Primeros pasos en Tableau, crear las primeras vistas:

- Ver el promedio del precio por barrio
- Ver el promedio del precio por 'bed type'
- Ver el promedio del precio por barrio y bed type

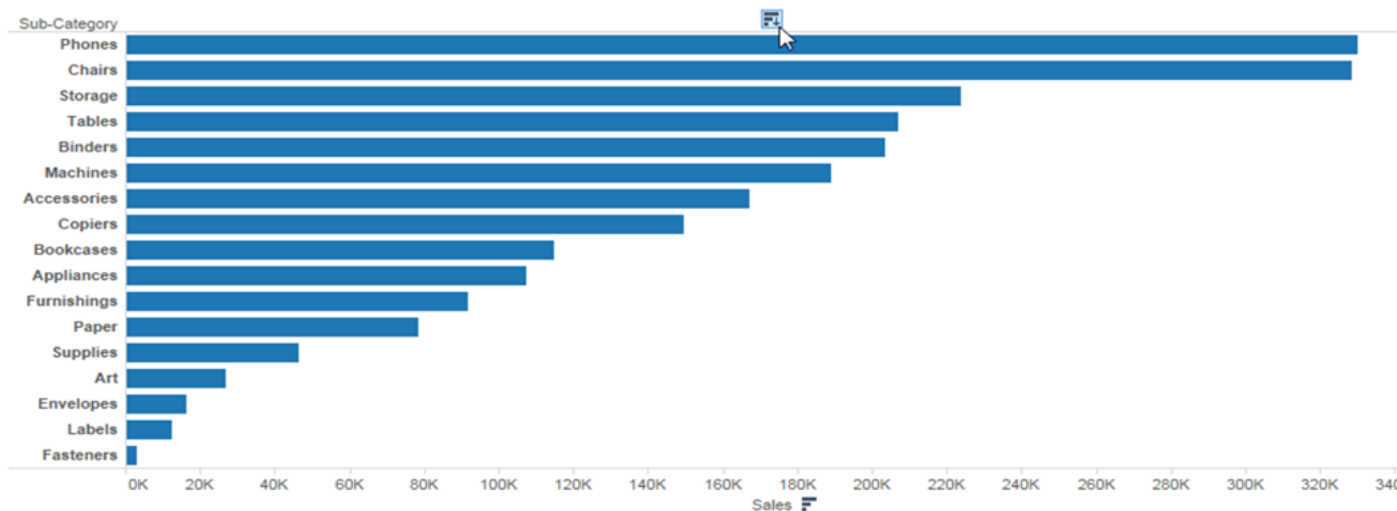
Ahora incluimos el factor tiempo

- Ver el promedio del precio por año.
- Ver el promedio del precio por mes.
- Ver el promedio del precio por mes para un año concreto.

# 3. Simplificación y ordenación de datos

## Ordenación de los datos

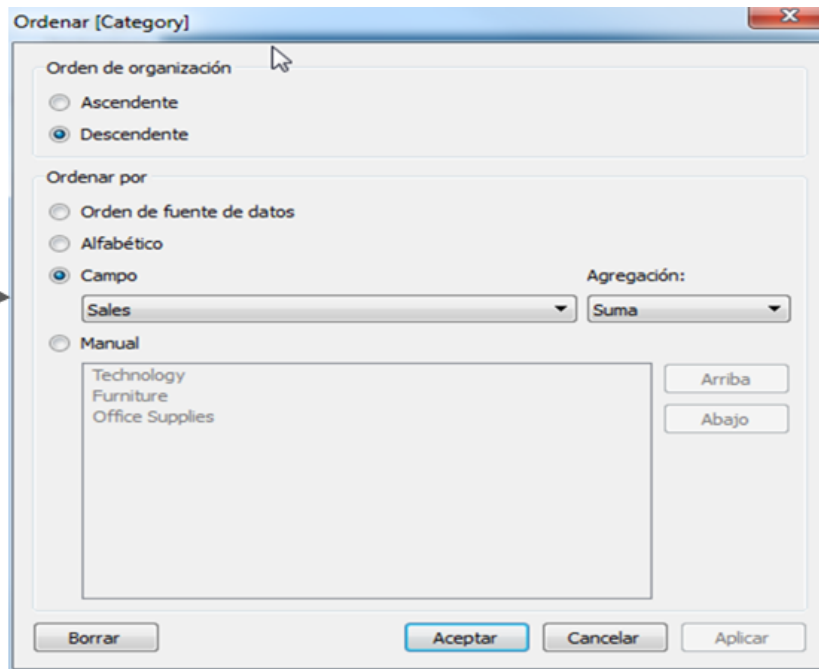
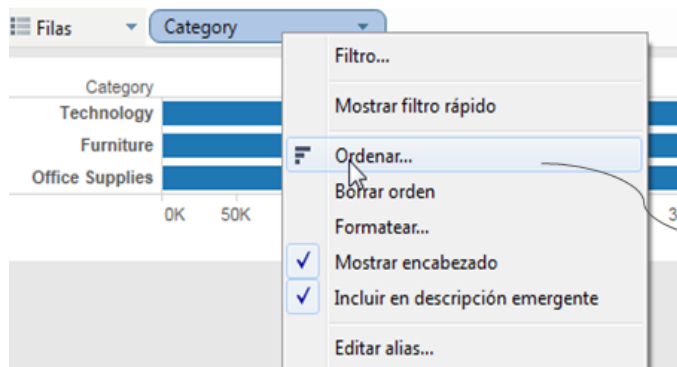
- Ordenación manual



### 3. Simplificación y ordenación de datos

#### Ordenación de los datos

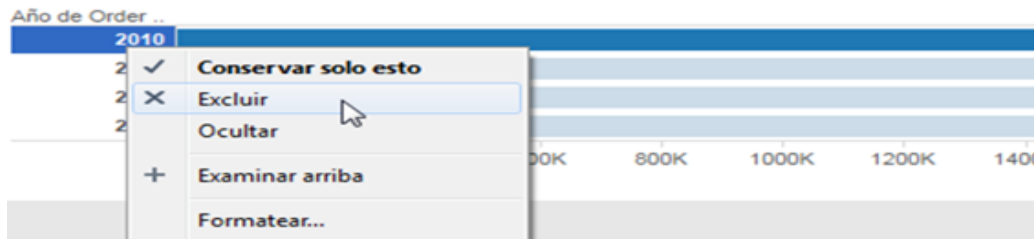
- Ordenación computerizada



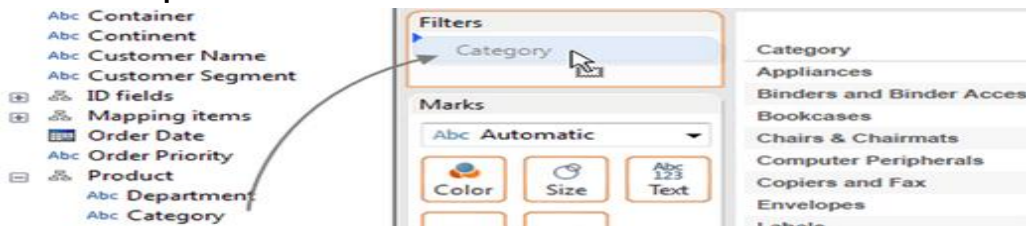
### 3. Simplificación y ordenación de datos

#### Filtros

- Permiten excluir/incluir determinados valores de una dimensión o métrica.
- Para incluir filtros en una vista:
  - Seleccionar un campo y elegir “Conservar solo esto” o “Excluir”



- Arrastrar un campo al estante filtros



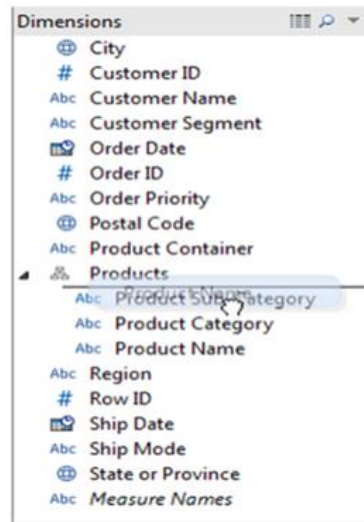
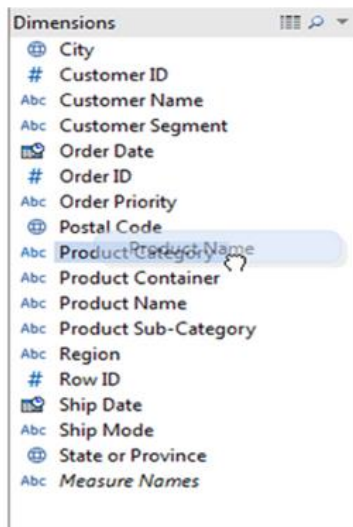
# Índice

1. Introducción a Tableau
2. Conexión a datos
3. Simplificación y ordenación de los datos
- 4. Organización de los datos**
5. División de datos por fecha
6. Mapeo geográfico de datos.
7. Panel de análisis
8. Personalización de los datos
9. Cálculos de tablas rápidos
10. Configuración y diseño avanzado de dashboards
11. Visualización de relación entre valores numéricos
12. Visualización de valores específicos
13. Visualización de un desglose de un todo
14. Visualización de distribuciones
15. Análisis visual

## 4. Organización de los datos

### Jerarquías

- Permiten al usuario navegar por los datos. Por ejemplo, una jerarquía llamada geografía que contenga País - Provincia - Ciudad.
- Para crear la jerarquía simplemente se arrastran unas dimensiones sobre otras.





# Ejercicios

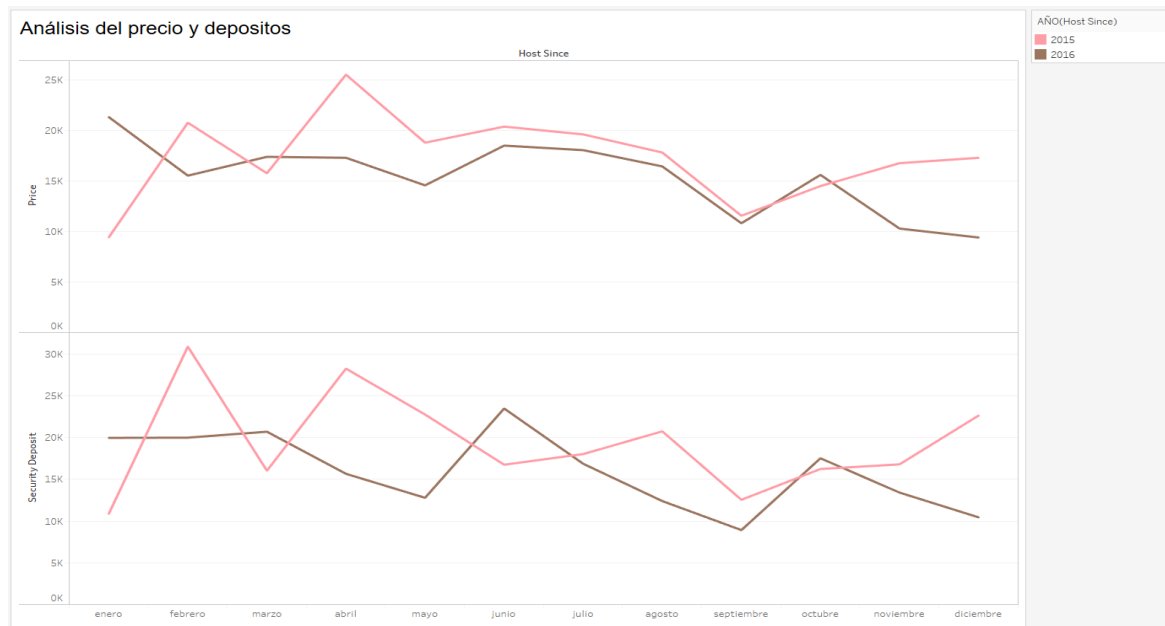


Crear las siguientes jerarquías:

- Jerarquía de tipo de estancia (Property type-Room type-Bed type)
- Jerarquía geográfica (country-city-zipcode)

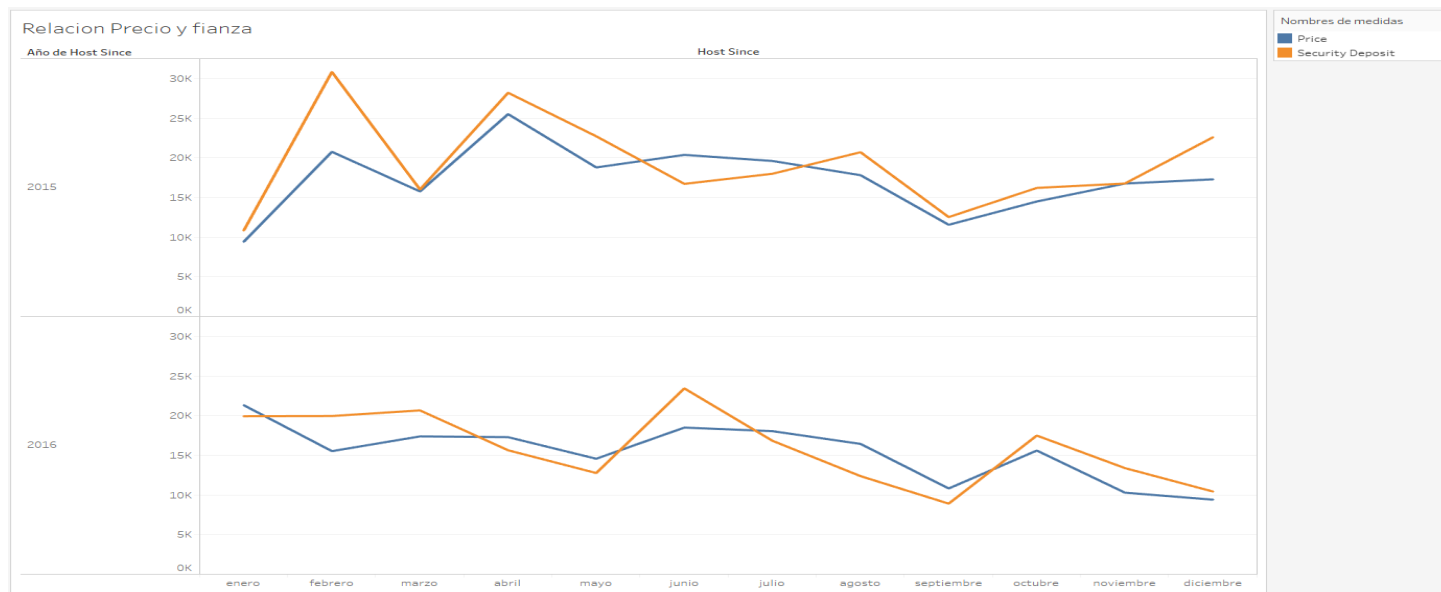
# Ejercicio 1

Análisis del precio y depósitos. ¿Cómo han evolucionado los precios y fianzas de manera mensual a lo largo de los 2 últimos años? Queremos representar esta información de manera que permita contrastar los precios y fianzas con los del año anterior.



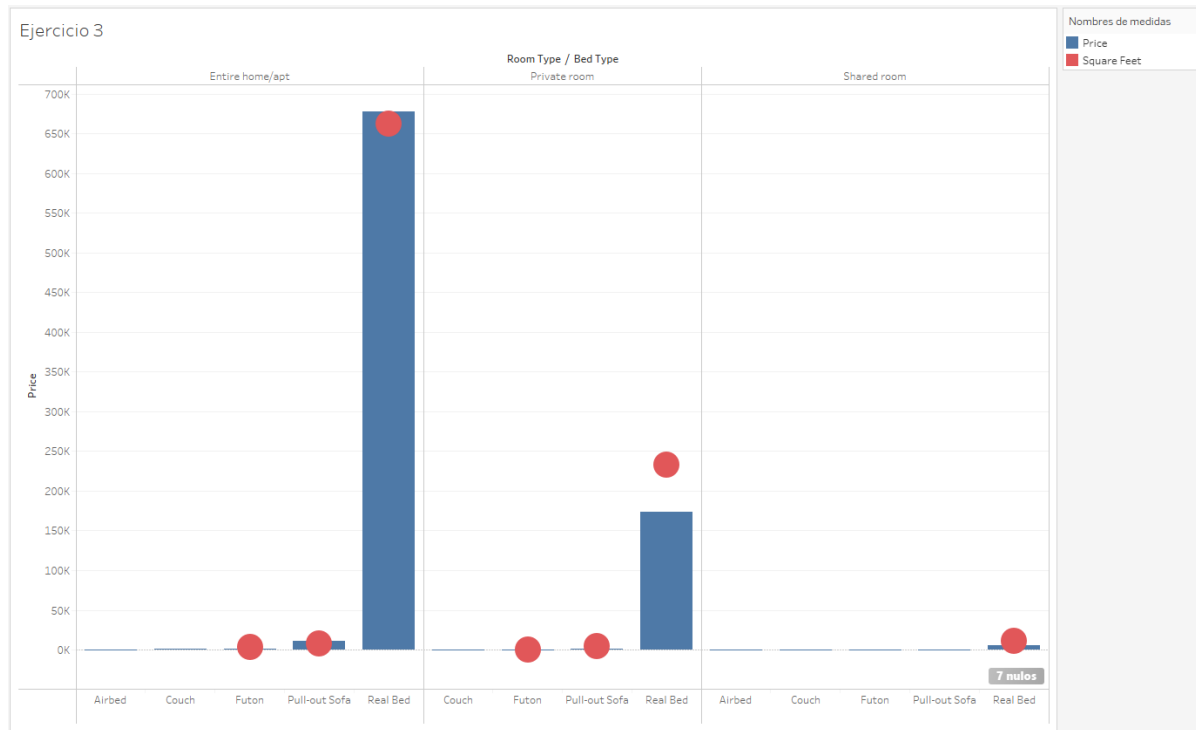
## Ejercicio 2

Construir una vista que muestre claramente cuál es la relación que existe entre los precios y las fianzas a lo largo del año. Mostrar esta relación para 2015 y 2016.



# Ejercicio 3

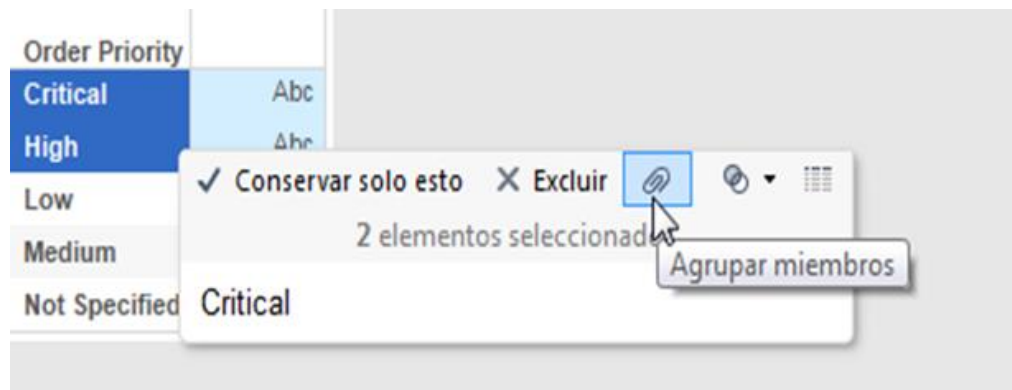
¿Cuál es la relación entre el precio y el tamaño de las viviendas?



## 4. Organización de los datos

### Grupos

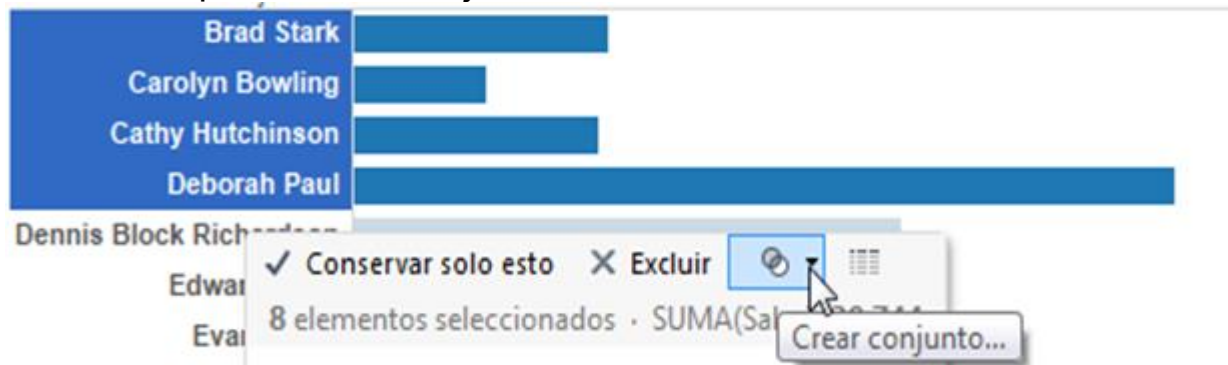
- Permiten agrupar varios miembros de una dimensión en categorías que no existen en los datos. Por ejemplo, agrupar usuarios según su tipología: usuario, supervisor, etc.
- La nueva dimensión que se crea se utiliza de la manera habitual en las vistas pero no es posible utilizarla para construir campos calculados.



## 4. Organización de los datos

### Conjuntos

- Subconjuntos de datos que cumplen unas ciertas condiciones. Por ej. clientes a los que se aplica un cierto descuento, zonas con unos gastos de envío determinados.
- Hay dos tipos de conjuntos: constantes y calculados
  - Constantes: sus miembros son fijos. Se crean seleccionando los elementos que se desea incluir y utilizando la opción “crear conjunto”.



## 4. Organización de los datos

### Conjuntos (II)

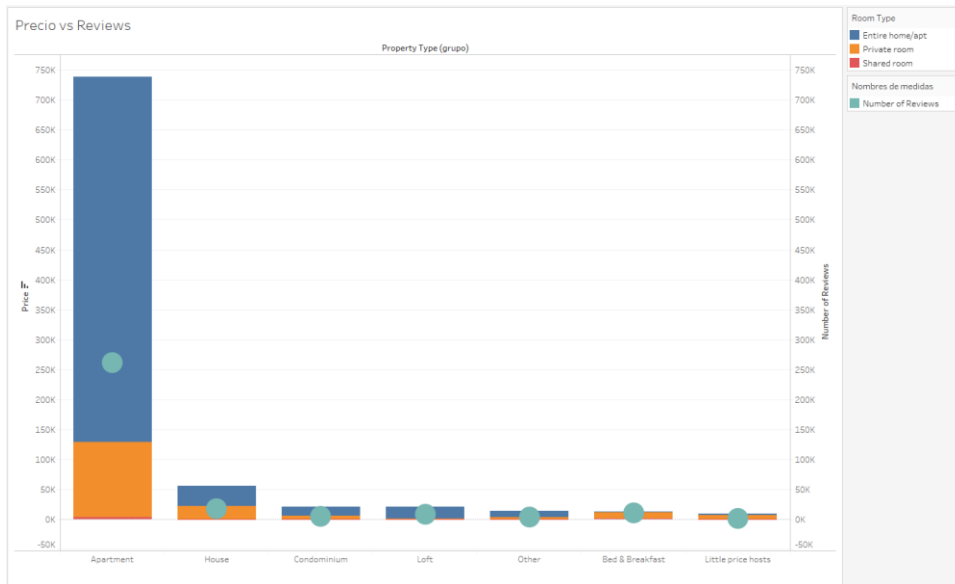
- Calculados: sus miembros se calculan de manera dinámica a partir de una fórmula o condición.

- Es posible crear combinaciones de conjuntos ya existentes.

# Ejercicio 4

A partir del ejercicio anterior, modificar el análisis para:

- Agrupar las tipos de vivienda que tienen menos de 20 reviews en una única categoría.
- Mostrar los precios por tipo de habitación.

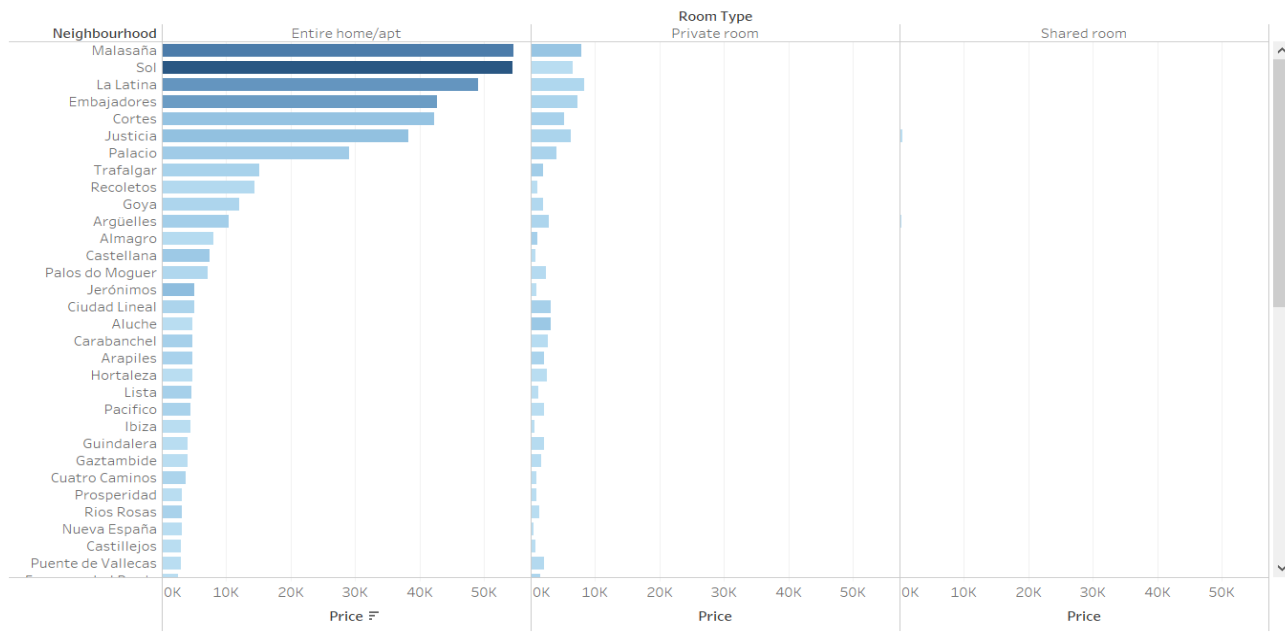




# Ejercicio 5

Realizar una visualización que muestre el precio y metros cuadrados por barrio y tipo de habitación. Además, debe existir la posibilidad de navegar tipo de habitación y tipo de cama.

Ejercicio 5



# Ejercicio 6

Construir un informe donde se muestre la tendencia del precio de cada barrio mes a mes.

