



# Microsoft Power BI

# Aviso

Se recuerda que esta sesión será grabada, salvo el listado de asistentes y las preguntas del chat, que no serán objeto de grabación. Si no has prestado tu consentimiento, recuerda desactivar tu cámara y micro, ya que si realizas una participación verbal esta constituirá tu consentimiento para la grabación

# Marcadores (bookmarks)

---

Partimos ahora de [esta versión](#) del ejemplo anterior

Vamos a ver distintas posibilidades de los marcadores (bookmarks)

Idea: los bookmarks capturan un estado concreto del informe de forma que podemos volver a este estado cuando deseemos



# Aplicación 1: estado de filtros

A screenshot of a filter panel from a web application. It contains six dropdown menus: 'Región' (with checkboxes for Europa, Japón, Norteamérica, and Otros), 'Año', 'Editorial', 'Plataforma', 'Género', and 'Nombre'. Each dropdown currently shows 'Todas'. To the right of the dropdowns is a small icon of a document with a magnifying glass.

Como en el panel hay muchos filtros, ir borrándolos todos cada vez que queremos volver al estado inicial puede ser muy molesto

Vamos a utilizar un marcador para “recordar” el estado inicial de los filtros

Pasos

- 1) Borramos a mano todos los filtros
- 2) En Ver + Marcadores se nos abre el panel de marcadores. Allí añadimos uno con el nombre que queramos ([reset\\_filters](#) p.ej.)
- 3) Añadimos un botón que tenga como asociado ese marcador ([imagen del botón](#) en Formato+Botón+Estilo+Icono+Tipo personalizado)

También se pueden crear para recordar combinaciones de filtros determinadas

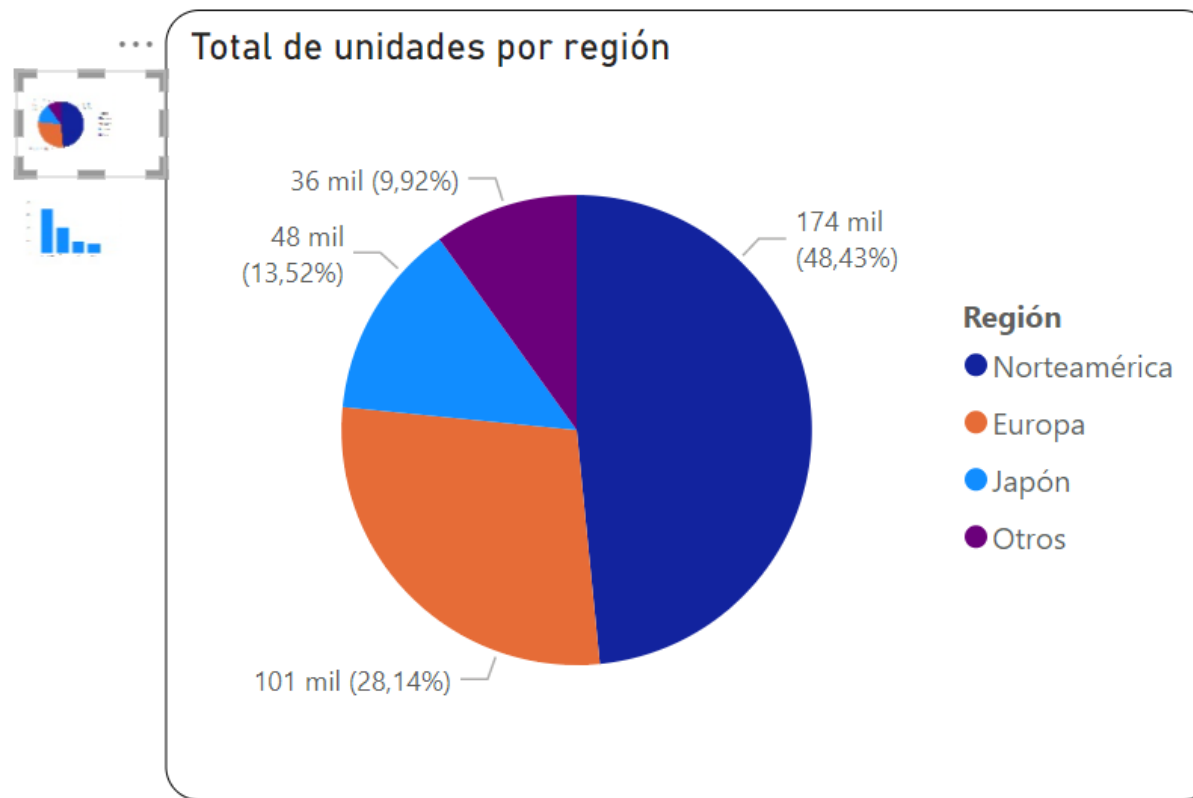
## Aplicación 2: alternar gráficos

Vamos a ver + Selección

Hacemos click en el gráfico que queramos y en el panel de selección le ponemos un nombre que podamos recordar

Ahora sobre el gráfico hacemos Ctr+C+Ctrl+V para crear un duplicado

Cambiamos el duplicado al tipo que queramos



# Aplicación 2

## alternar gráficos

- Jugando con los símbolos de ocultar podemos lograr que se vea un gráfico u otro
- A cada posibilidad le asociamos un filtro, con el nombre del gráfico que no se ve (ya que sirve para hacer que se vea)
- Añadimos dos botones con los iconos de [tarta](#) y [barra](#)

The screenshot displays a configuration panel for a dashboard, divided into two main sections: 'Selección' (Selection) and 'Marcadores' (Markers).

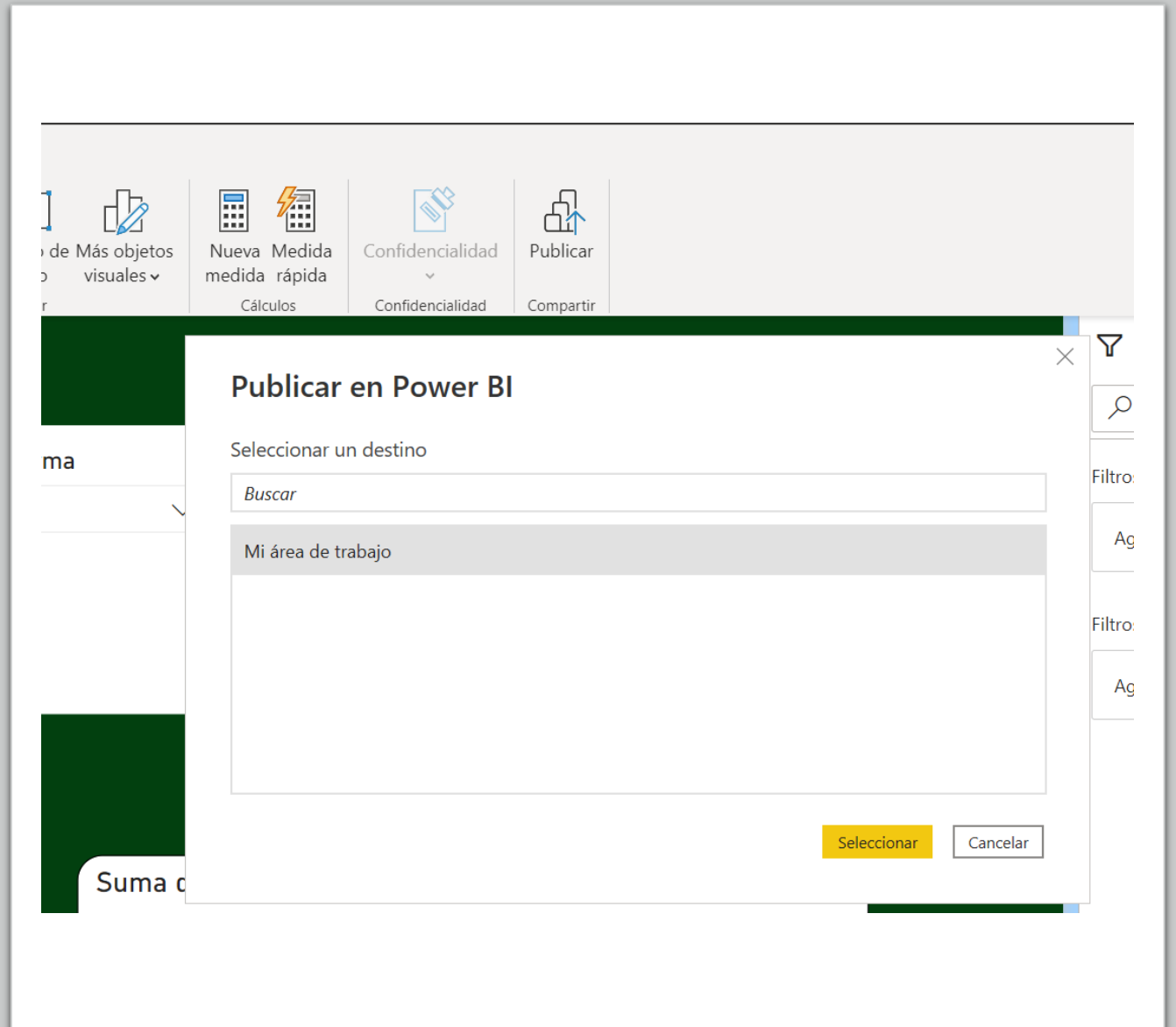
**Selección:** This section contains a table with columns for 'Orden de las capas' (Layer Order), 'Orden de tabulación' (Tabulation Order), and 'Mostrar / Ocultar' (Show / Hide). The 'Mostrar' column has a yellow bar above it. The 'Ocultar' column has a 'Mostrar' button and an 'Ocultar' button. The table lists various visualizations, including 'Botón', 'Cuadro de texto', 'Segmentación de datos', 'Suma de Unidades Vendidas por Año y Región', 'Suma de Unidades Vendidas por Región (barras)', 'Suma de Unidades Vendidas por Región (tarta)', and 'Forma'. The 'Suma de Unidades Vendidas por Región (barras)' and 'Suma de Unidades Vendidas por Región (tarta)' rows are highlighted with a blue background.

**Marcadores:** This section contains a list of markers, including 'reset\_filters', 'barras', and 'pie chart'. The 'pie chart' marker is highlighted with a grey background.

Orden de las capas	Orden de tabulación	Mostrar	Ocultar
Botón			☞ ...
Botón			☞ ...
Botón			☞ ...
Cuadro de texto			☞ ...
Segmentación de datos			☞ ...
Segmentación de datos			☞ ...
Segmentación de datos			☞ ...
Segmentación de datos			☞ ...
Suma de Unidades Vendidas por Año y Región			☞ ...
Suma de Unidades Vendidas por Región (barras)			☞ ...
Suma de Unidades Vendidas por Región (tarta)			☞ ...
Segmentación de datos			☞ ...
Segmentación de datos			☞ ...
Forma			☞ ...

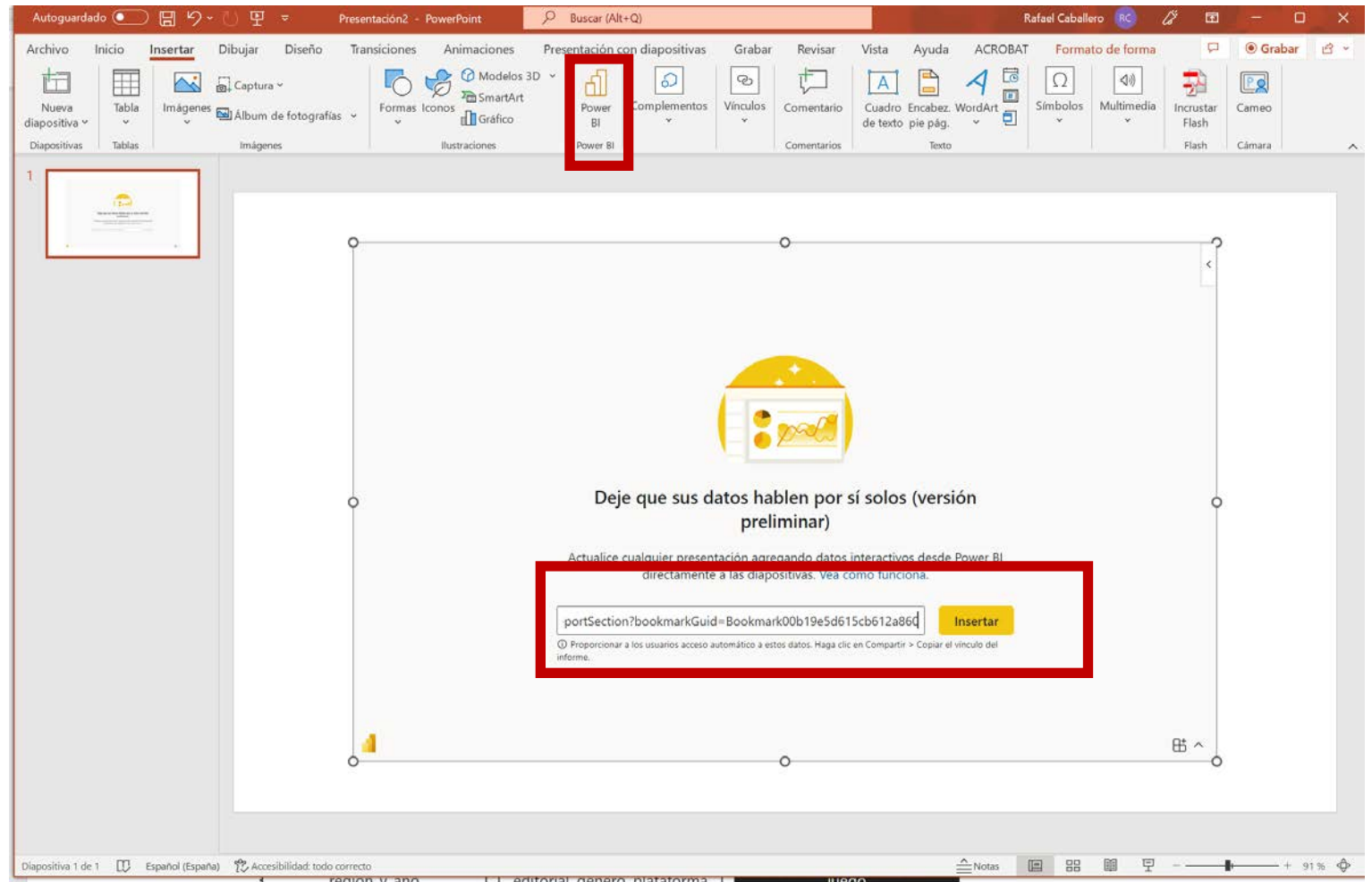
# Publicar el informe

Con el Inicio + publicar podemos publicar el informe en Power BI services (la nube)



# Incrustar en Power Point

---





# ¡Tenemos el informe en power point!

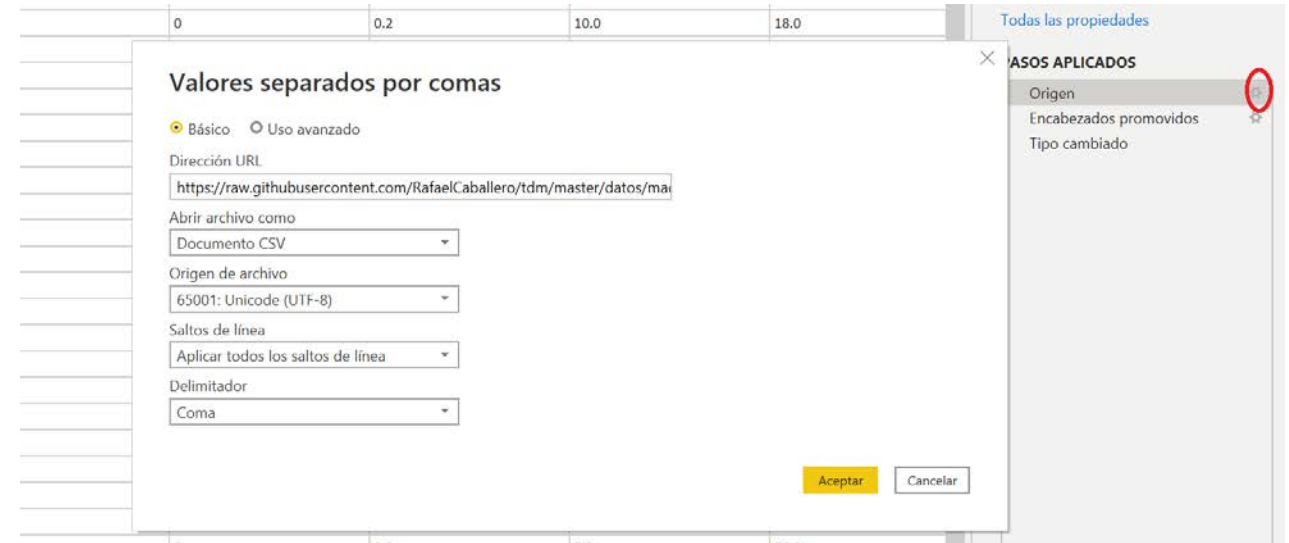


# Cambiar origen de datos

Fácil

Si es un cambio por ejemplo de un archivo local a otro archivo local, o de una URL a otra URL

- 1 Vamos a query editor
- 2 Seleccionamos la configuración de origen
- 3 Cambiamos los valores



# Cambiar origen de datos (II)

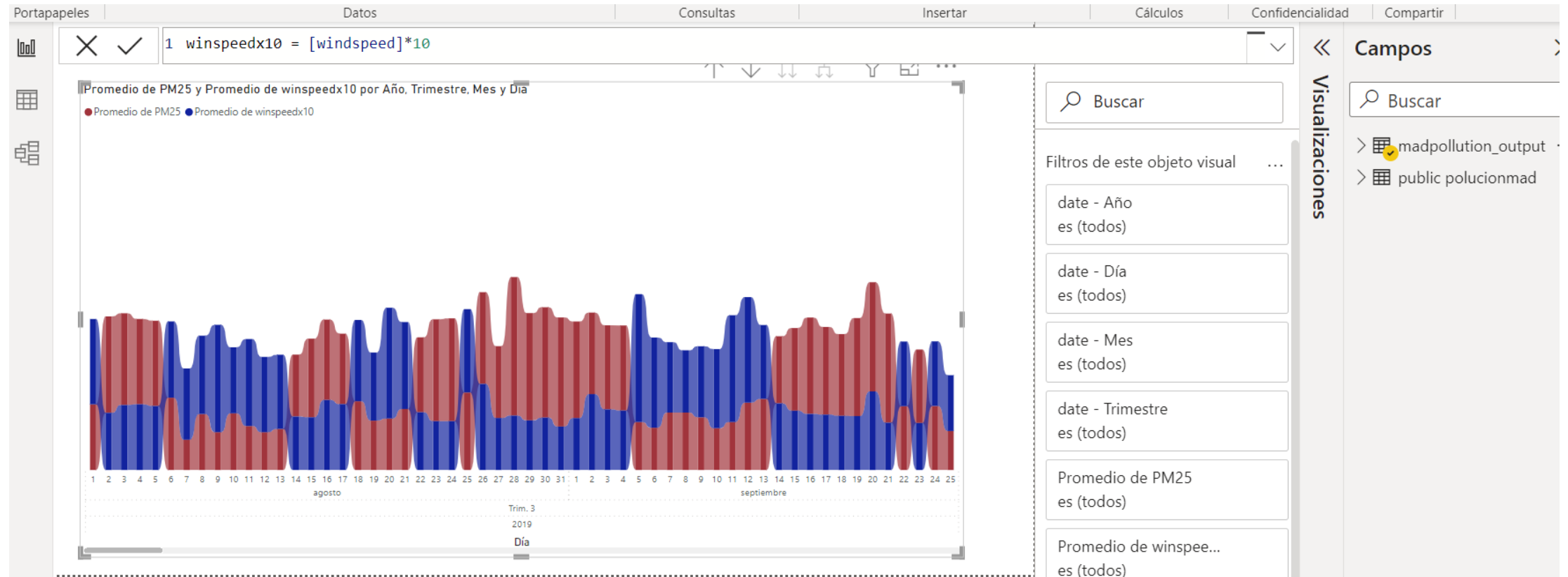
Caso más complejo: queremos pasar de una fuente de datos de tipo A a otra de tipo B.

Idea:

- Cargar la(s) tabla(s) de tipo B
- Editar el proceso de A y ver qué partes se pueden copiar (si son todas incluso se podría sustituir simplemente el proceso de carga de A por el de B)

# Ejemplo

Tenemos datos meteorológicos en CSV, pero queremos cambiarlos para que se lean de postgresQL



# Fichero de carga para la nueva fuente

```
let
  Origen = PostgreSQL.Database("127.0.0.1:5432", "postgres"),
  public_polucionmad = Origen{[Schema="public",Item="polucionmad"]}[Data]
in
  public_polucionmad
```

```
let
  Origen = Csv.Document(Web.Contents("https://raw.githubusercontent.com/RafaelCaballero/tdm/master/datos/madpollution_output.csv"),[Delimiter=";", Columns=21, Encoding=65001, QuoteStyle=QuoteStyle.None]),
  #"Encabezados promovidos" = Table.PromoteHeaders(Origen, [PromoteAllScalars=true]),
  #"Valor reemplazado" = Table.ReplaceValue(#"Encabezados promovidos", ".", ";", Replacer.ReplaceText,{"CO", "PM25", "NOx", "O3", "windspeed", "winddirection", "temperature", "humidity", "pressure", "rain", "traffic"}),
  #"Tipo cambiado" = Table.TransformColumnTypes(#"Valor reemplazado",{{"CO", type number}, {"PM25", Int64.Type}, {"NOx", Int64.Type}, {"O3", type number}, {"windspeed", type number}, {"winddirection", Int64.Type}, {"
  #"Columnas combinadas" = Table.CombineColumns(#"Tipo cambiado",{"year", "month", "day"},Combiner.CombineTextByDelimiter("/", QuoteStyle.None),"date"),
  #"Tipo cambiado1" = Table.TransformColumnTypes(#"Columnas combinadas",{{"date", type date}, {"hour", Int64.Type}, {"minute", Int64.Type}, {"second", Int64.Type}})
in
  #"Tipo cambiado1"
```

# Capítulo 4

## Modelado en PowerBI

1. Introducción al modelado de datos
  2. Administración de las relaciones de datos
  3. Creación de columnas calculadas
  4. Creación de medidas calculadas
- 

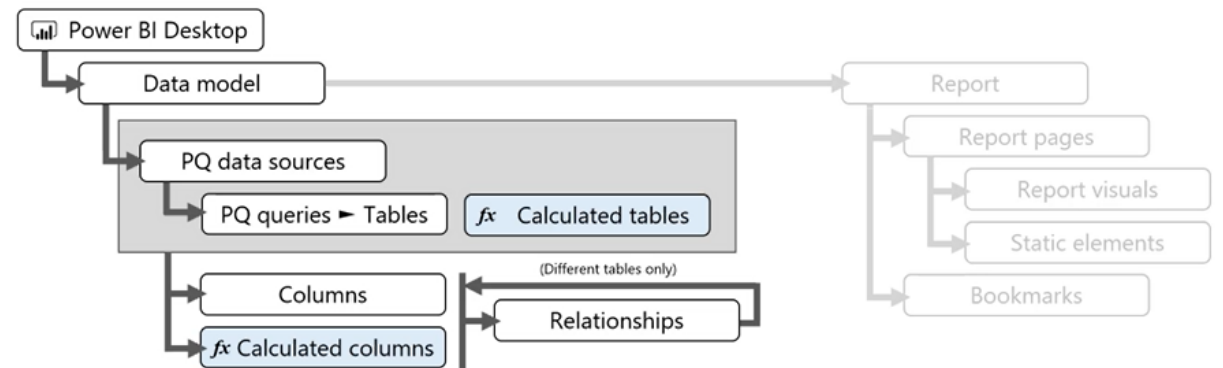


# Modelización en Power BI

Antes de hacer el report está la fase de modelización

En ella se establecen relaciones entre columnas

Estas relaciones son vitales para que los filtros en el report funcionen como es requerido



<https://learn.microsoft.com/es-mx/training/modules/dax-power-bi-models/1-introduction>

# Modelización ¿dónde estamos?

Bases de datos: 2 visiones

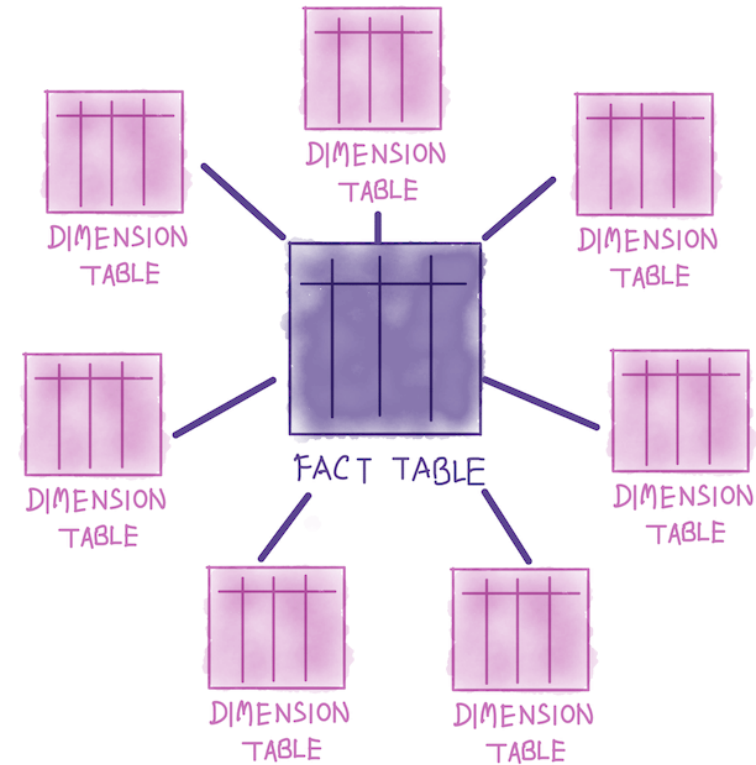


	Operacional	Analítica
Propósito	Desarrollo del negocio	Medida del desarrollo del neg.
Interacción principal	Insert,Update,Query>Delete	Query
Ámbito principal	Transacción individual	Agregaciones
Predictibilidad	Predecible y estable	Impredecible, cambiante
Optimización	Concurrencia, updates	Consultas eficientes
Principios diseño	Entidad Relación, 3FN	Dimensional (en estrella o cubos)
Términos relacionados	OLTP	Data Warehousre, Data Mart



# ¿Por qué crear un modelo dimensional?

- Más fácil realizar informes
- Consultas DAX más simples
- Más flexible y fácil de modificar



# Base Modelo Dimensional

- 2 conceptos hechos y dimensiones
- Ejemplo
  - ¿Cuál es la media del valor BrentOpen por meses?
- Van a depender de las consultas (informes)
  - hechos → el valor que se agrupa, que se muestra en distintas formas
  - Dimensiones → el contexto

# Hechos y dimensiones

- Los **hechos** son valores, a menudo numéricos, que se van a agrupar con respecto a otros
- Las **dimensiones** son los valores que se usan para agrupar los hechos a menudo mediante cierta **operación**

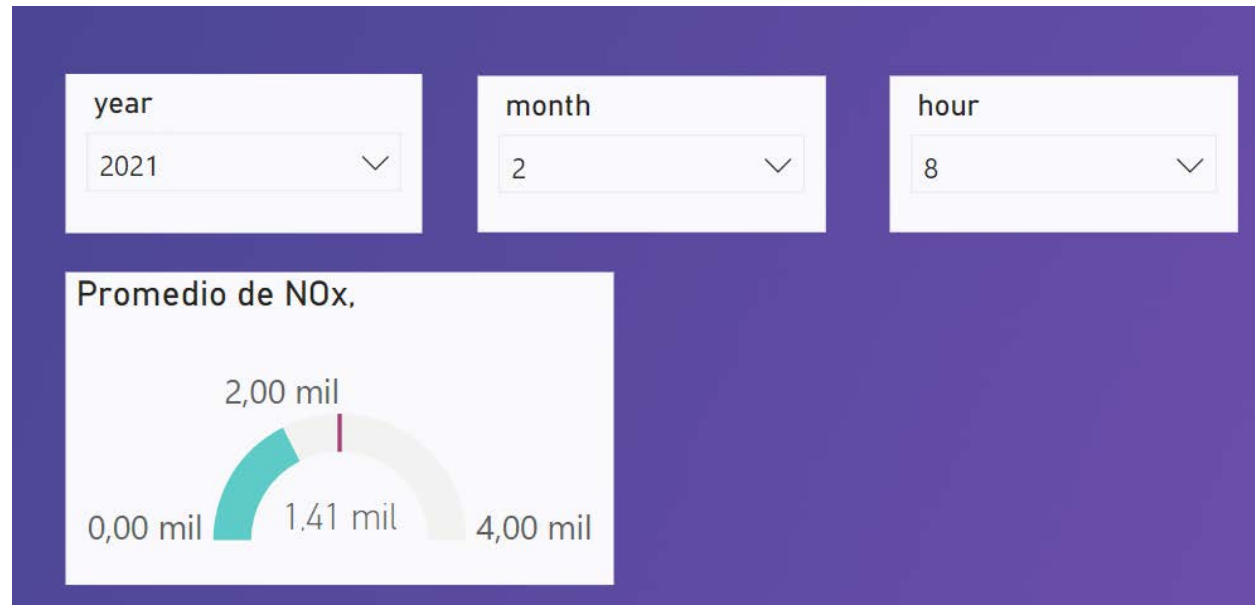
En un diagrama de barras las **dimensiones** serían las X y los **hechos** serían lo que se muestra en la Y agrupados por alguna **operación**

También son **dimensiones** todo aquello para lo que ponemos un filtro

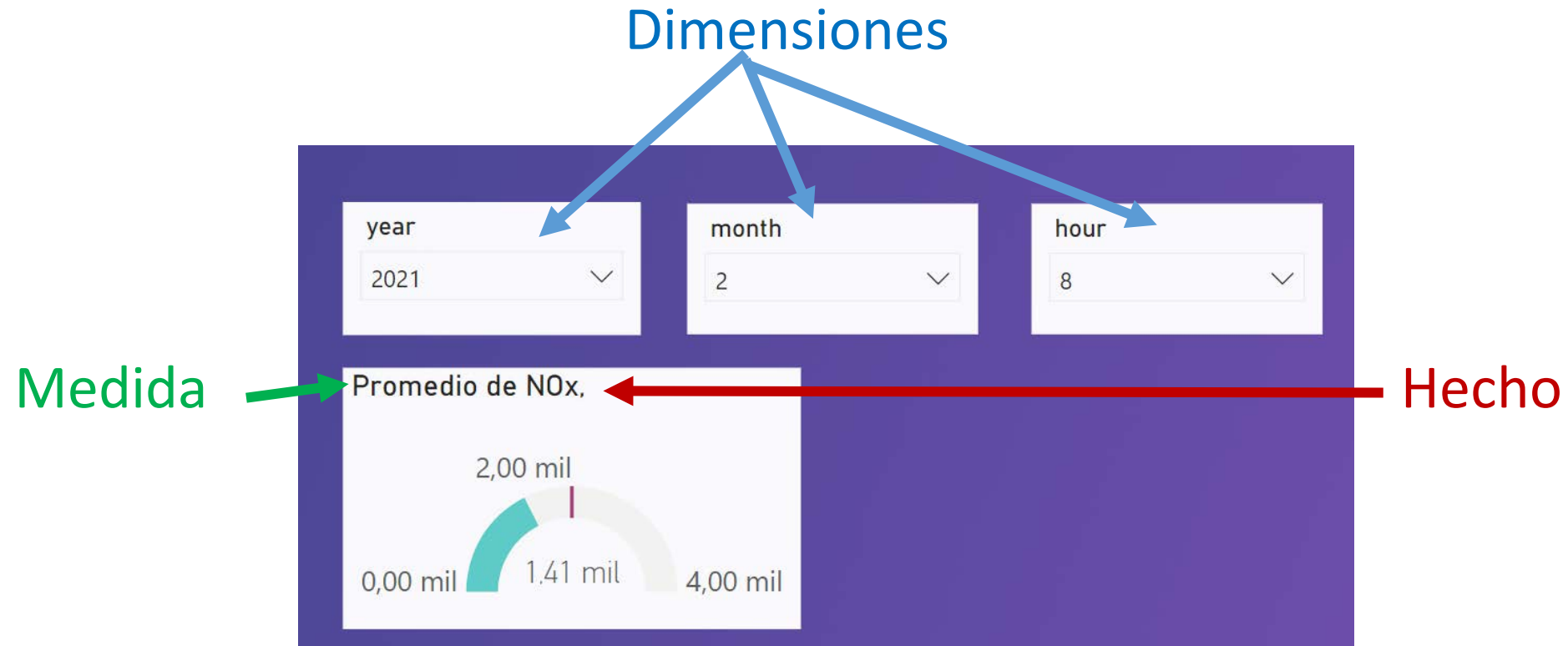
# Hechos y dimensiones

Ejemplo: obtener el **total** de **unidades vendidas** **por** cada **juego**

Ejemplo:

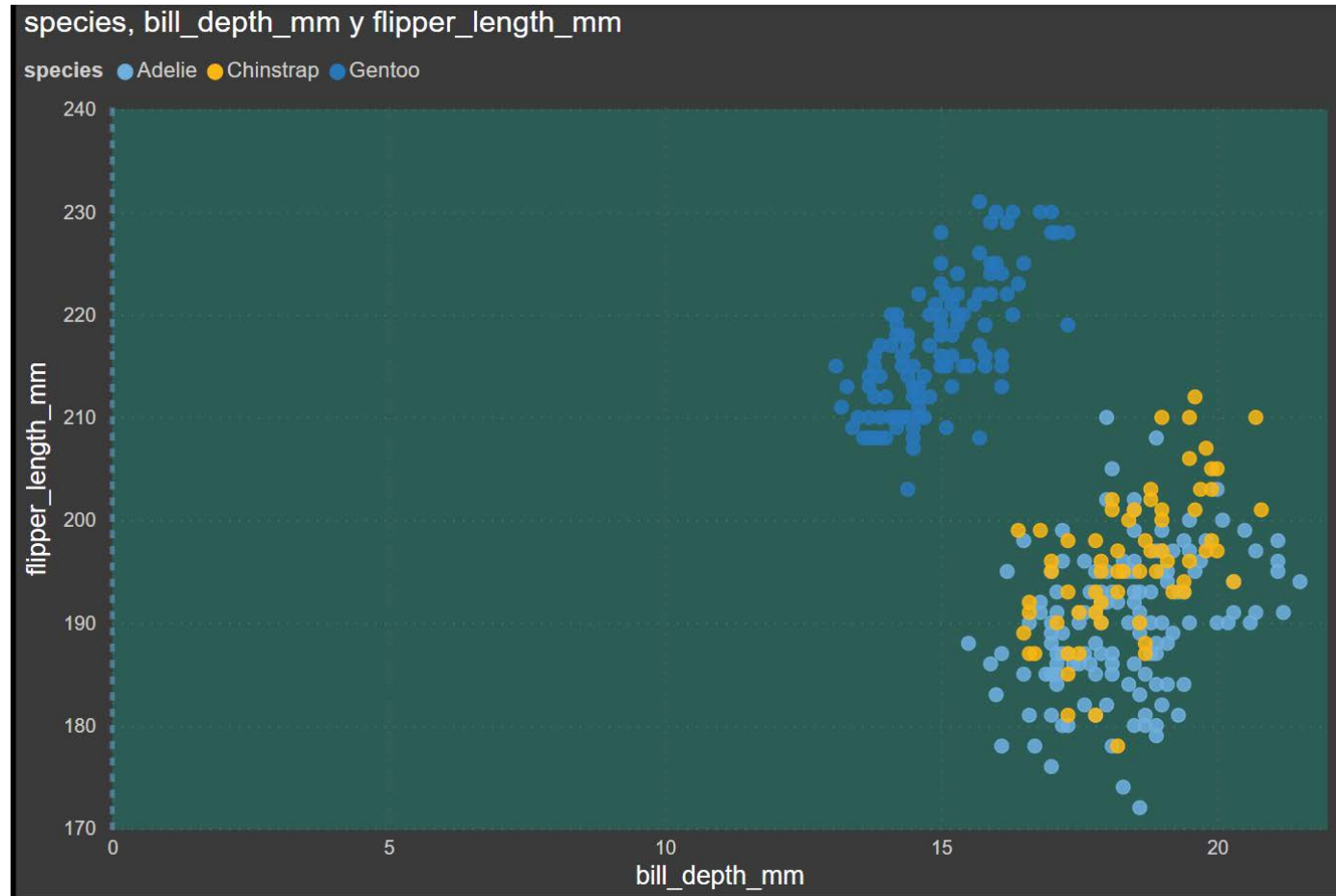


# Ejemplo (resuelto)



# Ejemplo: más difícil

¿Hechos y dimensiones?



***Dimensions provide the “who, what, where, when, why, and how” context surrounding a business process event. Dimension tables contain the descriptive attributes used by BI applications for filtering and grouping the facts.***

— Ralph Kimball, [The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling](#)

# Ejemplo

Hechos y dimensiones

## Facts

Order Dollars  
Cost Dollars  
Quantity Ordered

## Dimensions

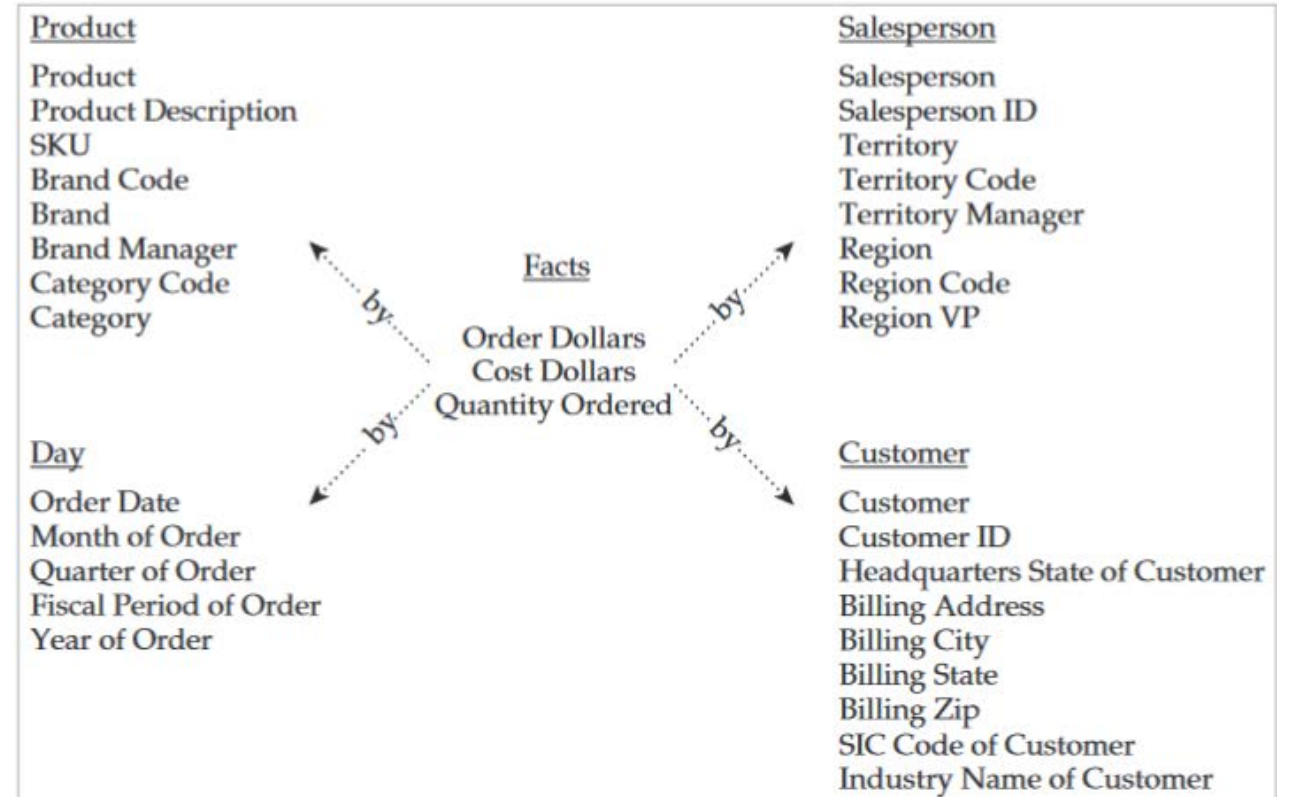
Product  
Product Description  
SKU  
Brand Code  
Brand  
Brand Manager  
Category Code  
Category  
Order Date  
Month of Order  
Quarter of Order  
Fiscal Period of Order  
Year of Order  
Salesperson  
Salesperson ID  
Territory  
Territory Code  
Territory Manager  
Region  
Region Code  
Region VP  
Customer  
Customer ID  
Headquarters State of Customer  
Billing Address  
Billing City  
Billing State  
Billing Zip  
SIC Code of Customer  
Industry Name of Customer





# Diagrama en estrella

- Una representación mediante tablas de los conceptos de hechos y dimensiones
- Los **hechos** se sitúan en el centro
- Alrededor las **dimensiones**, agrupadas en tablas que reúnen campos que se refieren a un concepto similar, proporcionan el contexto para analizar los hechos



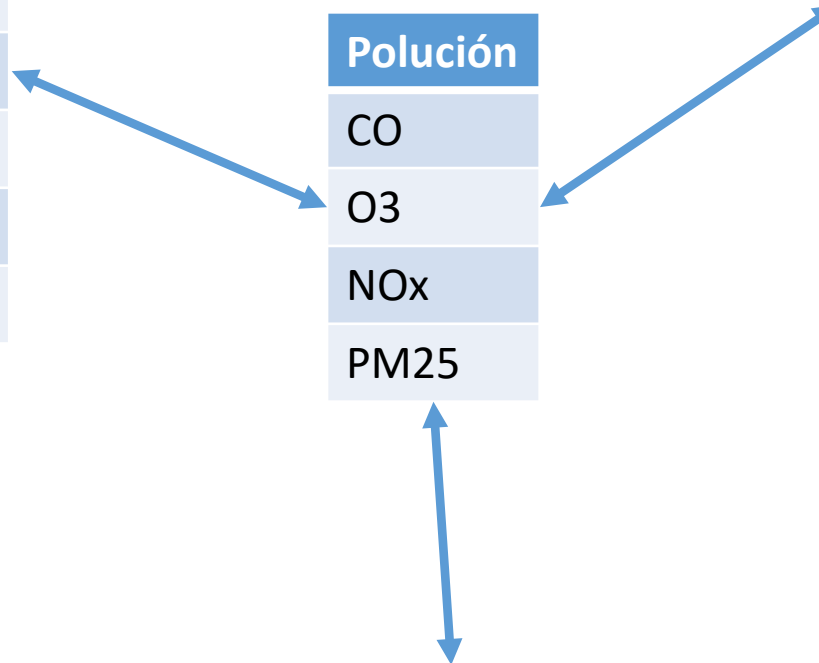
# Ejemplo

Clima
Temperature
Humidity
Pressure
Rain
Winddirection
Windspeed

Polución
CO
O3
NOx
PM25

Fecha
Year
Month
Day
Hour
Minute
laborday
Holiday
Sunday
saturday

tráfico
Traffic



✓ madpollution\_output

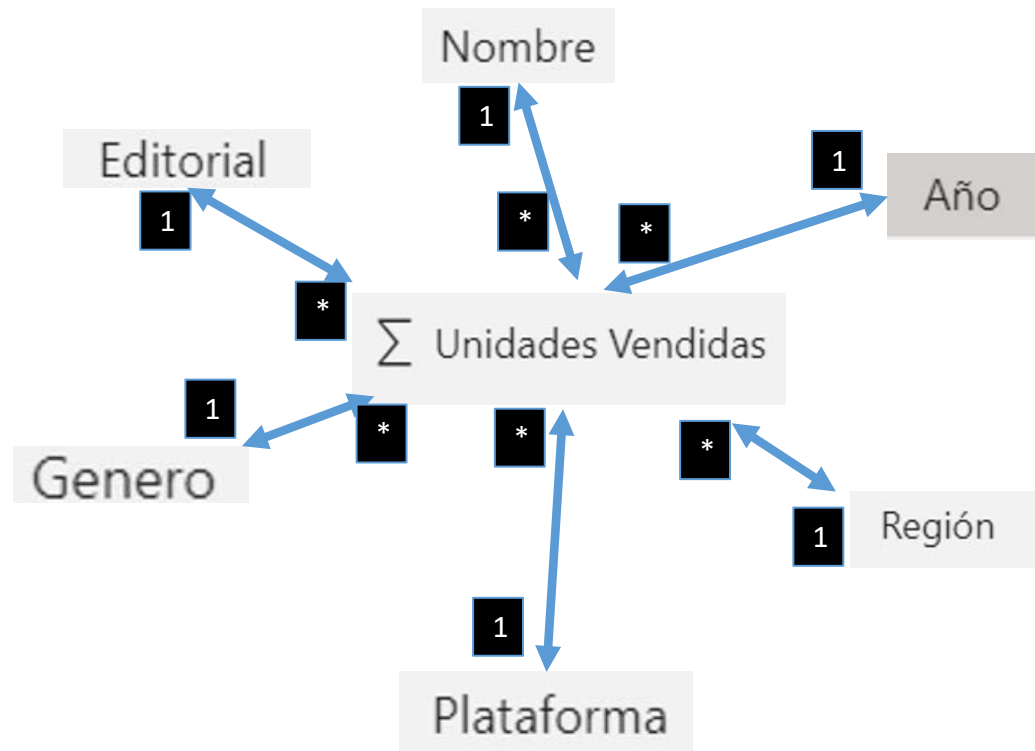
- ☐  $\Sigma$  CO
- ☐  $\Sigma$  day
- ☐  $\Sigma$  holiday
- ☐  $\Sigma$  hour
- ☐  $\Sigma$  humidity
- ☐  $\Sigma$  laborday
- ☐  $\Sigma$  minute
- ☐  $\Sigma$  month
- ☐  $\Sigma$  NOx
- ☐  $\Sigma$  O3
- ☐  $\Sigma$  PM25
- ☐  $\Sigma$  pressure
- ☐  $\Sigma$  rain
- ☐  $\Sigma$  saturday
- ☐  $\Sigma$  second
- ☐  $\Sigma$  sunday
- ☐  $\Sigma$  temperature
- ☐  $\Sigma$  traffic
- ☐  $\Sigma$  winddirection
- ☐  $\Sigma$  windspeed
- ☐  $\Sigma$  year

# Diagrama en Estrella: características

- Las relaciones entre las “puntas” y el centro son en general de “uno a muchos” o de “1 a 1”
- Las tablas a menudo no están en 3FN
- Si las tablas de las puntas a su vez se descomponen en otra y dan lugar a modelos “copo de nieve”



# Ejemplo



✓  videojuegos

Año

Editorial

Género

Nombre

Plataforma

Región

$\Sigma$  Unidades Vendidas

# Implementación

Clima
Clima_key
Temperature
Humidity
Pressure
Rain
Winddirection
Windspeed

Polución
Clima_key
Fecha_key
Traffic_key
CO
O3
NOx
PM25

Fecha
Fecha_key
Year
Month
Day
Hour
Minute
laborday
Holiday
Sunday
saturday

tráfico
Traffic_key
Traffic

✓ madpollution\_output

- ☐  $\Sigma$  CO
- ☐  $\Sigma$  day
- ☐  $\Sigma$  holiday
- ☐  $\Sigma$  hour
- ☐  $\Sigma$  humidity
- ☐  $\Sigma$  laborday
- ☐  $\Sigma$  minute
- ☐  $\Sigma$  month
- ☐  $\Sigma$  NOx
- ☐  $\Sigma$  O3
- ☐  $\Sigma$  PM25
- ☐  $\Sigma$  pressure
- ☐  $\Sigma$  rain
- ☐  $\Sigma$  saturday
- ☐  $\Sigma$  second
- ☐  $\Sigma$  sunday
- ☐  $\Sigma$  temperature
- ☐  $\Sigma$  traffic
- ☐  $\Sigma$  winddirection
- ☐  $\Sigma$  windspeed
- ☐  $\Sigma$  year

# Claves subrogadas

- Son claves, normalmente **númericas autoincrementales**
- Sirven para:
  - Identificar cada fila de una table dimensional
  - Asociarla a sus hechos relacionados
- Generalmente no tienen ningún sentido específico para el negocio

# ¿Cómo hacerlo en Power BI?

En transformación de datos:

- 0) Realizar las transformaciones habituales (tipos, unpivot)
- 1) Duplicar la tabla tantas veces como puntas tenga la estrella
- 2) Renombramos las copias para tener los mismos nombres que el modelo (por convención: primera consulta llevará el nombre de los hechos, el resto de las dimensiones)

# ¿Cómo hacerlo en Power BI?

## 3) Para cada dimensión

- Seleccionar las columnas que forman parte de la dimensión
- Eliminar duplicados
- Si la tabla no tiene una columna clave deberemos:
  - + Crear un índice que haga de clave
  - + Ir a la tabla de hechos y mezclarla con la nueva  
(usando los campos de la dimensión que haga falta)  
para añadir la nueva clave en la tabla de hechos



# ¿Cómo hacerlo en Power BI?

4) En la tabla de hechos dejar solo los campos relativos a los hechos más las claves de cada una de las dimensiones

5) Cerrar y aplicar los cambios y crear las relaciones en la vista de modelo

# Ejemplo

Clima
Clima_key
Temperature
Humidity
Pressure
Rain
Winddirection
Windspeed

Polución
Clima_key
Fecha_key
Traffic_key
CO
O3
NOx
PM25

tráfico
Traffic_key
Traffic

Fecha
Fecha_key
Year
Month
Day
Hour
Minute
laborday
Holiday
Sunday
saturday

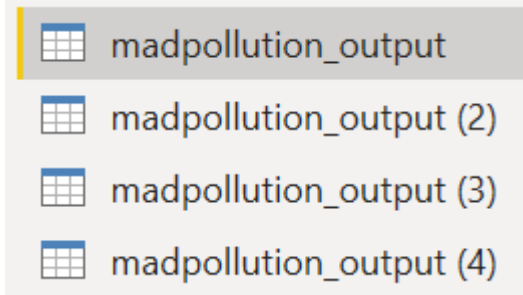
✓ madpollution\_output

- ☐  $\Sigma$  CO
- ☐  $\Sigma$  day
- ☐  $\Sigma$  holiday
- ☐  $\Sigma$  hour
- ☐  $\Sigma$  humidity
- ☐  $\Sigma$  laborday
- ☐  $\Sigma$  minute
- ☐  $\Sigma$  month
- ☐  $\Sigma$  NOx
- ☐  $\Sigma$  O3
- ☐  $\Sigma$  PM25
- ☐  $\Sigma$  pressure
- ☐  $\Sigma$  rain
- ☐  $\Sigma$  saturday
- ☐  $\Sigma$  second
- ☐  $\Sigma$  sunday
- ☐  $\Sigma$  temperature
- ☐  $\Sigma$  traffic
- ☐  $\Sigma$  winddirection
- ☐  $\Sigma$  windspeed
- ☐  $\Sigma$  year

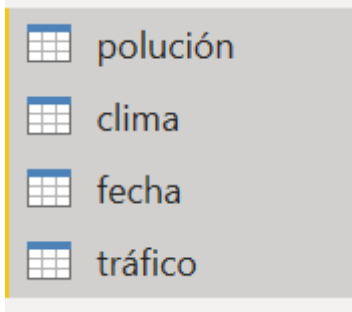
# Pasos

0) Comprobar tipos numéricos; crear “date”

1) Duplicamos 3 veces



2) Renombramos



# Pasos

3)

- Seleccionamos campos,
- Quitamos duplicados
- Creamos índice
- Combinamos con la tabla “polución”

## Combinar

Seleccione una tabla y las columnas coincidentes para crear una tabla combinada.

polución

PM25	NOx	O3	windspeed 1	winddirection 2	temperature 3	humidity 4	pressure 5	rain 6	traffic
10	29	58,87	1,84	97	26,1	52	943	0	570
10	18	63,73	1,97	117	24,9	55	943	0	404
9	19	66,5	1,72	96	24	55	943	0	287
10	15	66,62	1,55	106	23,3	55	943	0	209

clima

windspeed 1	winddirection 2	temperature 3	humidity 4	pressure 5	rain 6	clima_key
1,84	97	26,1	52	943	0	1
1,97	117	24,9	55	943	0	2
1,72	96	24	55	943	0	3
1,55	106	23,3	55	943	0	4
1,13	67	22,9	57	943	0	5

Tipo de combinación

Externa izquierda (todas de la primera, coincidencias...

☒ Use las coincidencias aproximadas para comparar la combinación.

► Opciones de coincidencia aproximada

✓ La selección coincide con 14749 de 14749 filas de la primera tabla.

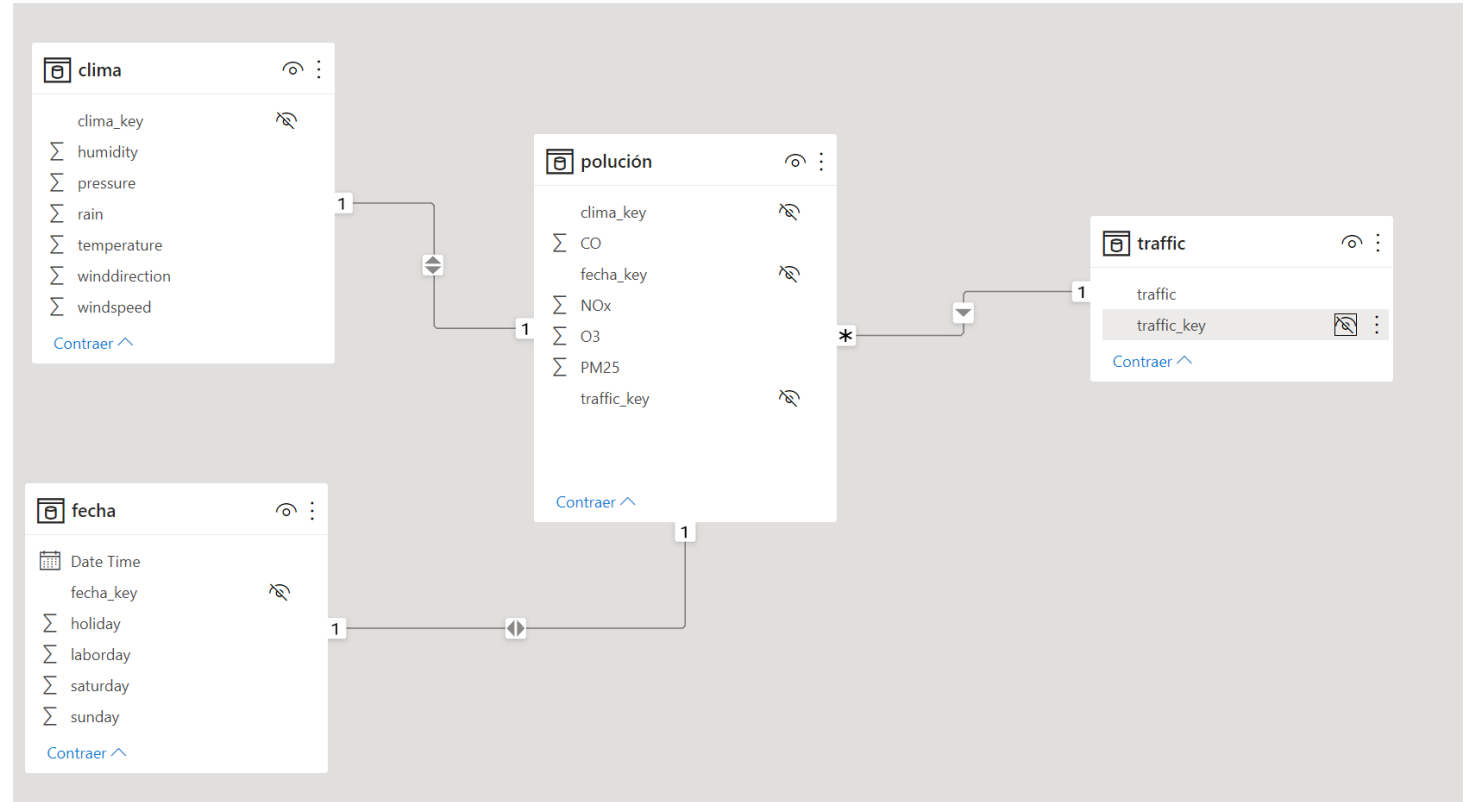
Aceptar

Cancelar

# Pasos

4) Eliminamos los campos de la tabla polución

5) Conectamos en la vista de modelo



# Ejercicio modelización

screen_name	text	localidad	hora	verified	friends_count	followers	tweets	Persona Verbal	valor persona	Tiempo Verbal	Valor Tiempo	renta	tipo_renta	sentiment
noeliaciado_c	Vivir a base de arroz, pavo, y pescado hervido se está convirtiendo en modo de vida 🍛🍷🍷 #lapupas	Madrid,Madrid,España	16	0	255	90	4151	tercera	1	presente	1	20650	<23000	0

[Este informe](#) incluye datos de tweets.

La fuente de datos incluye en cada fila datos de un tweet y de su usuario

**screen\_name**: id. en Twitter del usuario

**Text**: Texto del tweet

**Localidad**: dónde reside el usuario

**Hora**: hora a la que fue emitido el tweet

**Verified**: si es usuario verificado

**friends\_count**: nº. de us. a los que sigue

**Followers**: nº de seguidores del usuario

**Tweets**: Total de tweets emitidos por el usu.

**Persona\_verbal**: persona en la que está el verbo del tweet (si tiene varios verbos con diferentes personas se repetirá la fila)

**Valor persona**: número >0 de veces que de esa persona verba

**Tiempo verbal**: tiempo verbal del tweet (si tiene varios tiempos se repetirá la fila)

**Valor tiempo**: número >0 de veces que aparece ese tiempo verbal en el tweet

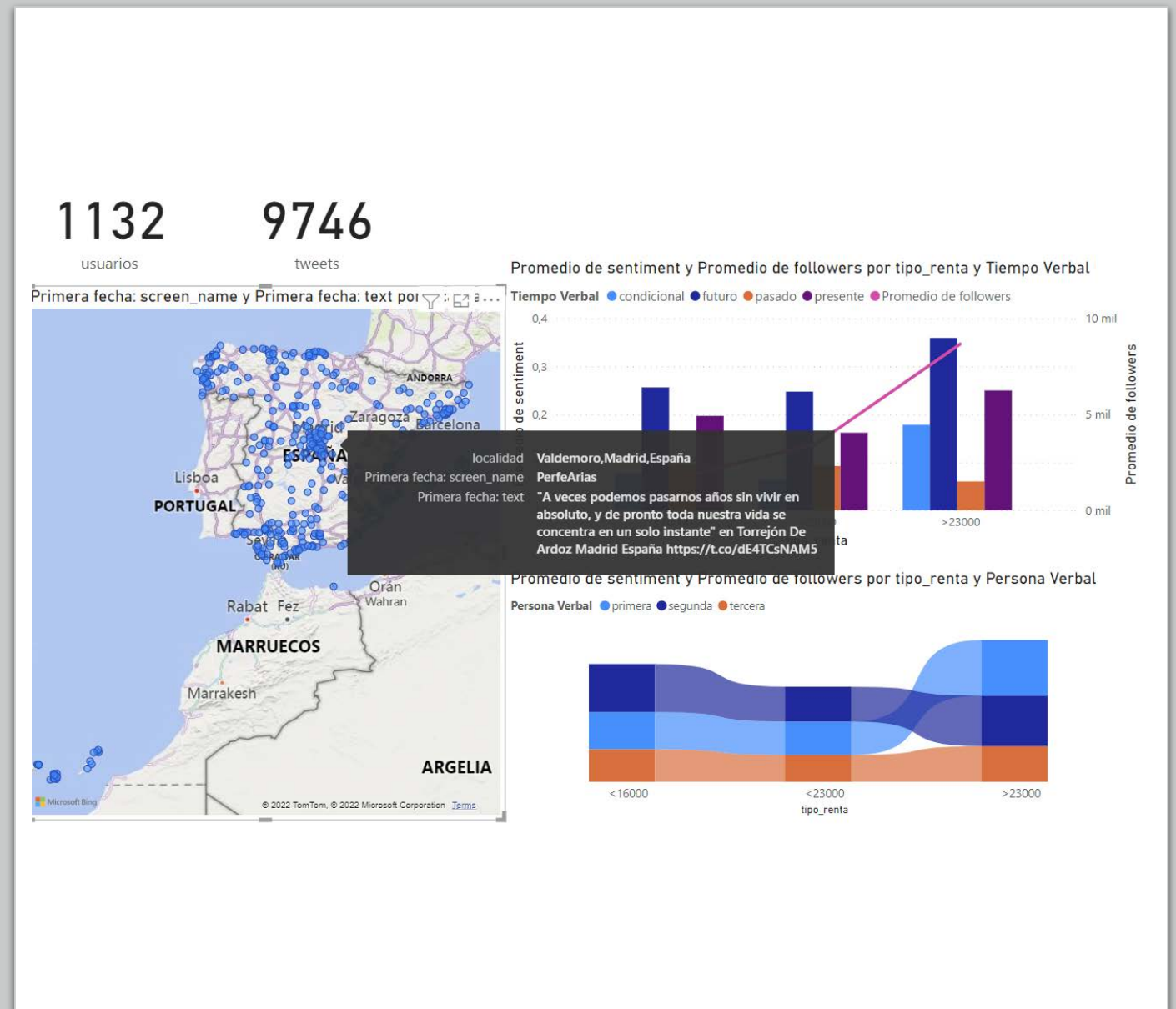
**Renta**: Renta media de la zona donde vive el usuario

**Tipo renta**: un string que divide la renta en 3 tipos (bajo, medio, alto)

**Sentiment**: sentimiento del tweet entre -1 (negativo) y positivo

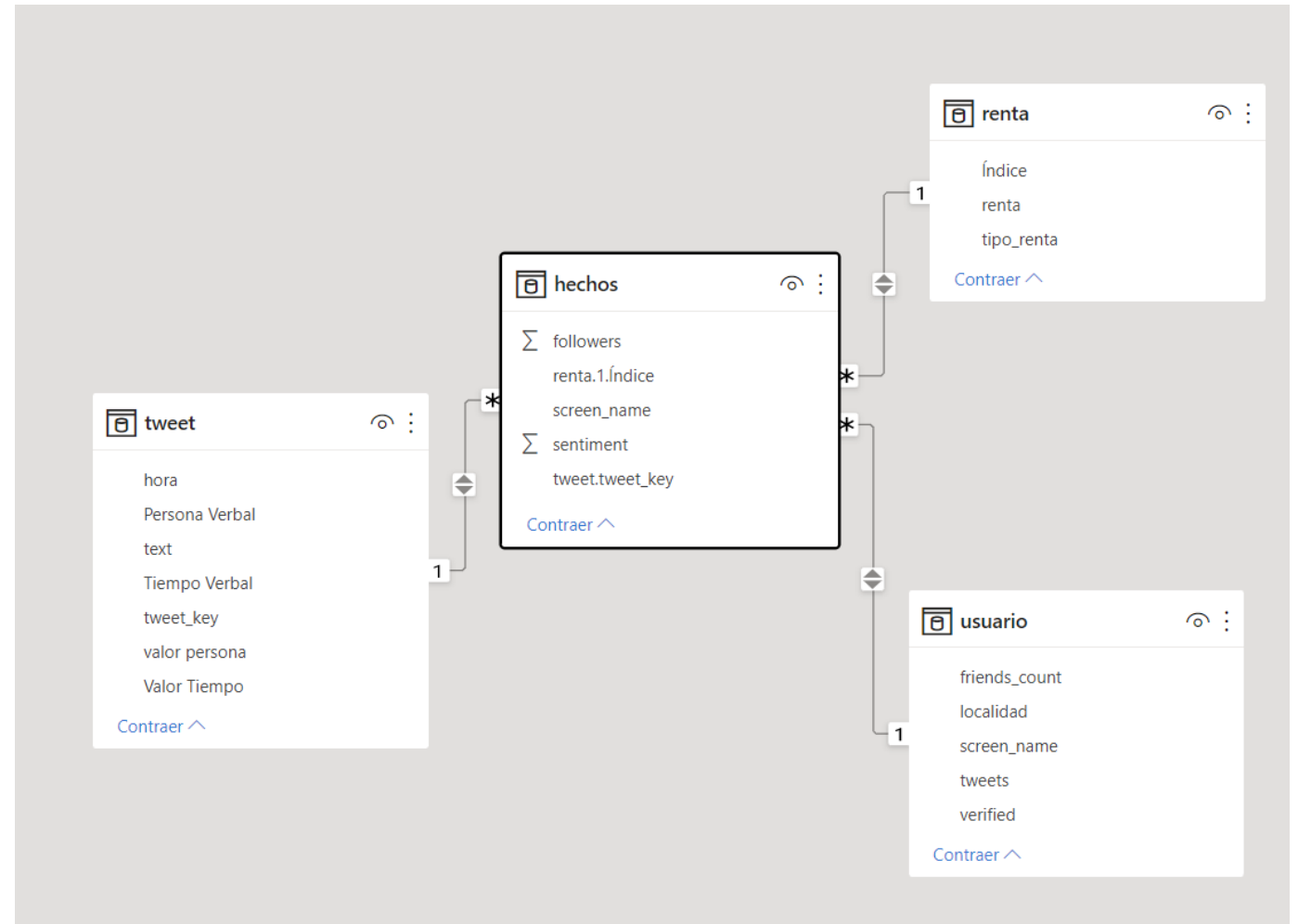
# Ejercicio modelización

Aplicar la modelización en diagrama en estrella considerando como hechos el **sentimiento**, y el número de **followers**



# Ejercicio

- Tiene sentido hacer una tabla de renta ya que no es la renta del usuario sino de una zona
- También tendría sentido hacer una tabla de localización, pero eso solo un campo → copo de nieve





# Algunas consideraciones finales

- ¡Modelos analíticos actualizables!
  - En principio no están pensados para ser actualizados (nuevas filas)
  - ...pero Power query ha registrado todos los pasos que hemos hecho y actualizará el modelo en estrella al actualizar los datos
- ¿Varias tablas de hechos?
  - Sí, es posible, pero deben ser las mismas dimensiones
  - Muy habitual cuando dentro de los hechos tenemos distinta granularidad: unos se recogen por meses, otros por días

# Bibliografía

- *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling*. Ralph Kimball, Margy Ross
- *Star Schema The Complete Reference*. Christopher Adamson
- <https://learn.microsoft.com/en-us/power-bi/guidance/star-schema>
- <https://www.linkedin.com/pulse/data-modeling-power-bi-common-beginner-mistakes-vitor-klein>
- <https://www.elegantbi.com/post/top10bestpractices>



Fechas en  
Power Query

# Generando las fechas: una utilidad


- Es muy habitual tener **dimensiones fecha**
- Suele ser un listado de fechas que enlazamos con la tabla de hechos para señalar que se puede filtrar por fecha
- Un problema habitual es **generar esta tabla**
  - Podemos extraer las fechas a partir de la tabla inicial, pero puede que haya **huecos**
  - También que nos falten valores, como el día de la semana
- Una posible solución es utilizar un script como el que se incluye en [esta página](#) o [este](#) un poco más completo

# Generando fechas: una utilidad

Para ejecutarlo, dentro de Power Query:

- Vamos a **inicio**, **nuevo origen**, **consulta en blanco**
- Ahora vamos a **Vista + Editor Avanzado**
- Borrar el código por defecto y copiar a cambio el script
- Nos pregunta por el rango de fechas

Escribir parámetros

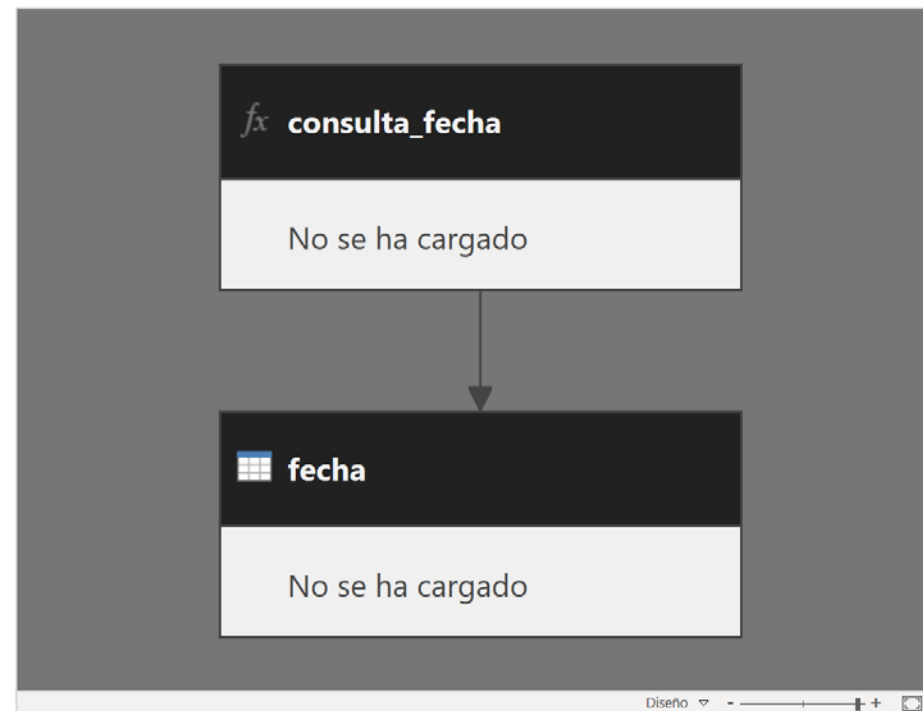
StartDate  
 

EndDate  
 

# Generando fechas: una utilidad

- El resultado es una función y la tabla que genera para los datos que le hemos dado a los *parámetros de entrada*
- Existe una dependencia entre ambas que podemos ver en **Vista + Dependencias de la consulta**

Dependencias de la consulta



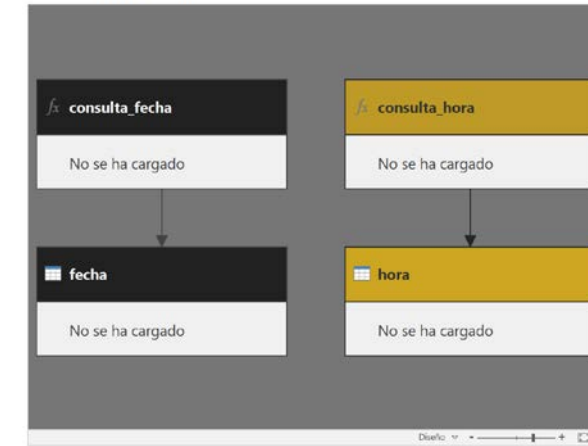
# Pero, ¿y las horas? ¿y los minutos? ¿y los segundos?

- Si se quiere tratar con horas, minutos o segundos lo mejor en general es tenerlo como algo **aparte de la fecha**, aunque siempre podemos combinar columnas para crear un objeto **DateTime**
- ¿Cómo incorporar la hora? Vamos a hacerlo en dos pasos
  - Generar una tabla con los números de 1 a 24 en power query
  - Hacer el producto cartesiano de ambas tablas

# Hora: generar tabla

- Podemos usar [este script](#)
- El funcionamiento es similar al del script anterior
  - Generamos una consulta en blanco
  - Vamos al editor avanzado
  - Copiamos y pegamos el código
- Nos pregunta por el valor inicial y final, así puede valer también para minutos y/o segundos

Dependencias de la consulta



## Escribir parámetros

NumeroInicial

NumeroFinal

Invocar

Borrar



# Añadiendo hora: combinar ambas

- No es necesario, podemos dejar “hora” como una pequeña tabla de look-up (sobre todo si se añade algún campo como por ejemplo “parte del día”, “mañana”, “tarde”, “noche”...
- Si queremos añadir la hora tendremos que hacer un *producto cartesiano* con la tabla de fechas
  - Seleccionar la tabla de fechas
  - Damos añadir columna personalizada
  - En el nombre de la columna podemos poder lo queramos
  - En la formula pondremos el nombre de la tabla con la fecha
  - Ya solo hay que expandir la columna generada

## Columna personalizada

Agregue una columna que se calcula a partir de otras columnas.

Nuevo nombre de columna

hora

Fórmula de columna personalizada ⓘ

= hora

[?] hora  
[?] consulta\_hora  
[?] HexagonSmartApi.ApplySelectList  
[?] HexagonSmartApi.ApplyUnitsOfMeasure  
[?] HexagonSmartApi.ExecuteParametricFilterOnFilterRecord  
[?] HexagonSmartApi.ExecuteParametricFilterOnFilterUrl  
[?] HexagonSmartApi.Feed  
[?] HexagonSmartApi.GenerateParametricFilterByFilterSourceType  
[?] HexagonSmartApi.GetODataMetadata  
[?] HexagonSmartApi.Typecast  
[?] SapHanaRangeOperator.Equals  
[?] SapHanaRangeOperator.GreaterThan

Columnas disponibles

Date

ter

< Number

th Number

Number

th

th

<< Insertar

✓ No se han detectado errores de sintaxis.

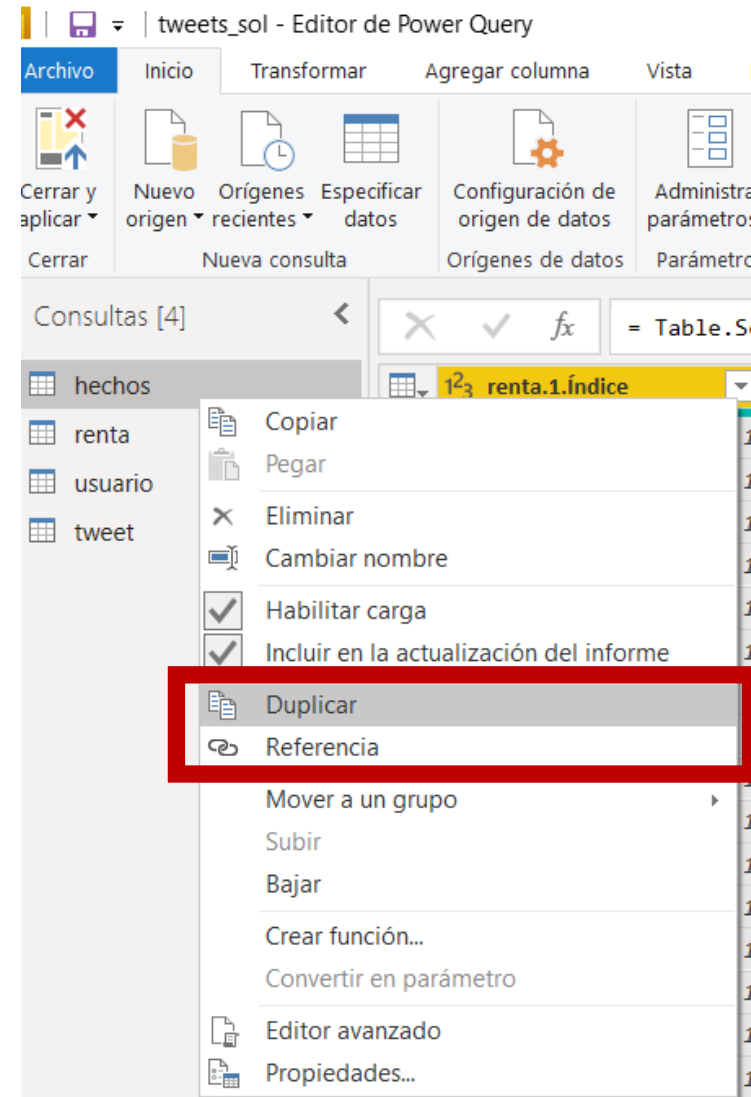
Aceptar

Cancelar

# Diferencia entre duplicar y referencia

Hay 2 opciones para copiar tablas (en realidad 3 incluyendo copiar y pegar)

- Duplicar
- Referencia



# Duplicado versus referencia

**Duplicado:** crea una copia de la tabla con todos sus pasos de preprocesado incluidos; se trata de una copia perfecta

Tras la copia las dos tablas son independientes, añadir pasos a la primera no afectará a la segunda

**Referencia:** la nueva tabla comienza con un solo paso, la toma de datos de la tabla original

Cualquier cambio que se haga en la tabla original afectará a la ref.

Interesante para crear “bifurcaciones” o pruebas de alternativas

```
let  
  Origen = tweet  
in  
  Origen
```

# Referencias: características

- No se pueden establecer **referencias circulares**, lo que impide combinar una tabla con otra que la referencie (Ej: combinación con tabla de hechos en el modelado en diagrama de estrella)
- Si una tabla da lugar a otra por referencia **no se puede** borrar sin haber borrado previamente aquella que la referencia
- Si se **copia y pega** una tabla creada por referencia se copiará y creará también el origen de la tabla

# Detección de referencias

Se puede ver las referencias con  
Vista + Dependencias de la consulta

Se distingue por la frase “no se ha  
cargado”

