

Sistemas multidimensionales

Práctica 2. Transformaciones con Power Query

Alumno: Alvaro Marín Pérez

A 23 de marzo de 2021

Contenido

[1. Importación de datos 3](#_Toc67497836)

[1.1. Eliminar filas 3](#_Toc67497837)

[1.2. Definir la cabecera en una sola línea 3](#_Toc67497838)

[1.3. Completar datos 4](#_Toc67497839)

[1.4. Eliminar dinamización de la tabla 4](#_Toc67497840)

[1.5 Definir las columnas de la tabla 4](#_Toc67497841)

[Ejercicio 3.1. 5](#_Toc67497842)

[1.5. Definir los datos a nivel de individuo 5](#_Toc67497843)

[1.6. Agrupar los datos a nivel de individuo 6](#_Toc67497844)

[1.7. Obtener medidas adicionales de los individuos 6](#_Toc67497845)

[Ejercicio 3.2 7](#_Toc67497846)

[1.8. Definir niveles adicionales de campos 8](#_Toc67497847)

[Ejercicio 3.3 10](#_Toc67497848)

[Ejercicio 3.4 11](#_Toc67497849)

[1.9. Generar hechos y dimensiones 11](#_Toc67497850)

[1.10. Generar las dimensiones 11](#_Toc67497851)

[Ejercicio 3.5 14](#_Toc67497852)

[Ejercicio 3.6 14](#_Toc67497853)

[3.5.2 Generar los hechos 15](#_Toc67497854)

[1.11. Seleccionar las tablas que se devuelven a Power BI 17](#_Toc67497855)

[Ejercicio 3.7 17](#_Toc67497856)

[Ejercicio 3.8 18](#_Toc67497857)

[2. Modelo y consultas en Power BI 19](#_Toc67497858)

[2.1. Definición de medidas calculadas 19](#_Toc67497859)

[Ejercicio 4.1 20](#_Toc67497860)

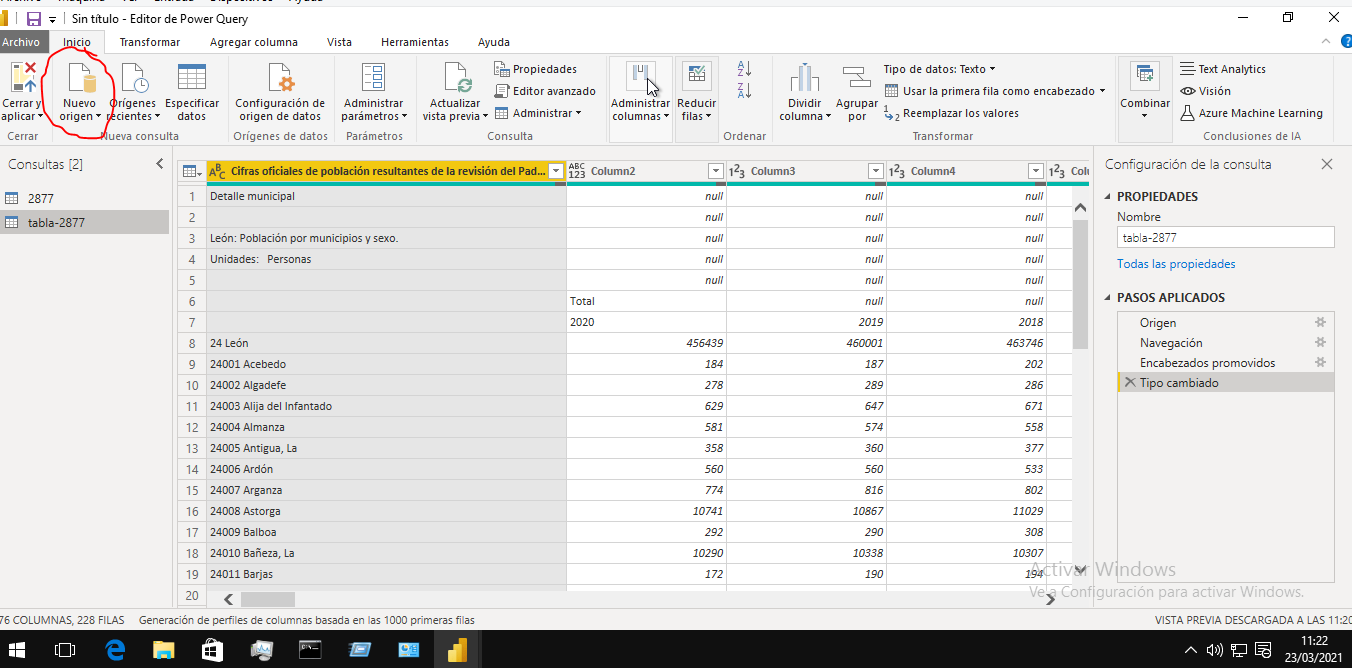
[2.2. Definición de jerarquías 21](#_Toc67497861)

[Ejercicio 4.2 22](#_Toc67497862)

[2.3. Consultas basadas en jerarquías 23](#_Toc67497863)

[Ejercicio 4.3 24](#_Toc67497864)

# Importación de datos



Seleccionamos el formato del archivo que tengamos(en nuestro caso Excel) y seleccionamos las hojas que importaremos.

## Eliminar filas

Podemos observar que no tenemos una tabla plana. Esto es debido a que los pasos aplicados (a la derecha de la pantalla) no son los correctos.

Para ellos, arreglaremos esto eliminando los pasos que no son los correctos, eliminando “Tipo cambiado” y “Encabezados promovidos”.

Si nos fijamos, se nos han incluido 6 filas con datos adicionales que a nosotros no nos sirven. Las eliminamos.

Inicio -> Reducir filas -> 6 filas

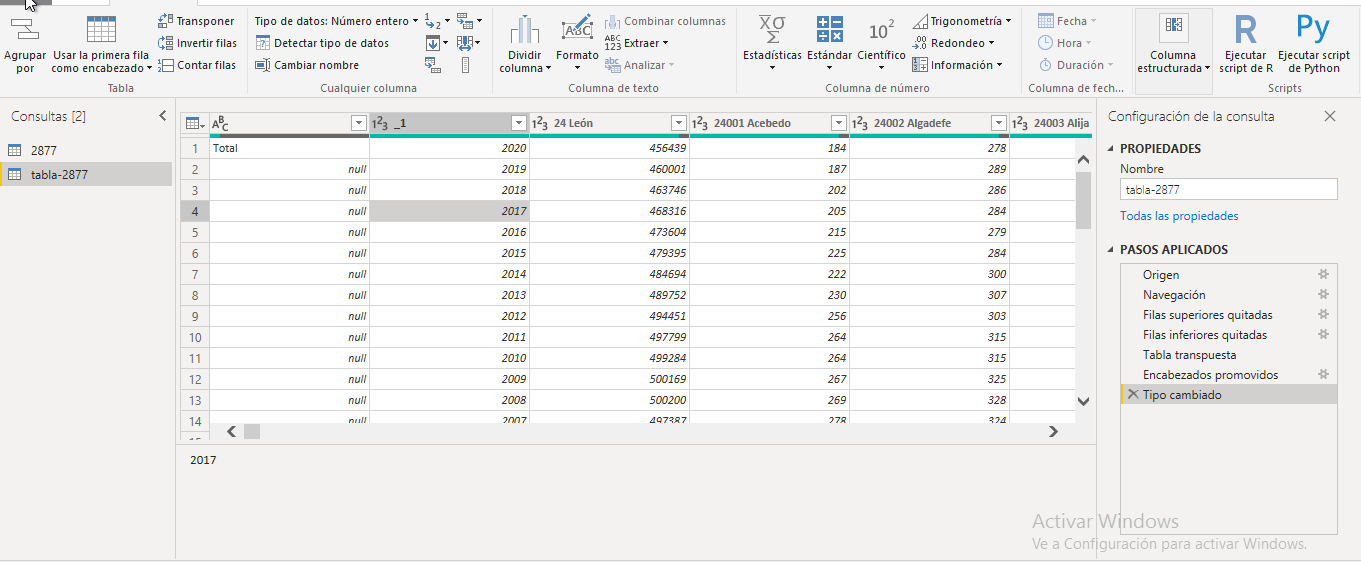
Hacemos lo mismo para las filas inferiores que también nos sobran.

## Definir la cabecera en una sola línea

Transformar -> Tabla -> usar la primera fila como encabezado

Hemos colocado la cabecera en una sola línea, en lo alto.

Eliminamos el paso “Tipo cambiado” porque, generalmente, siempre falla.

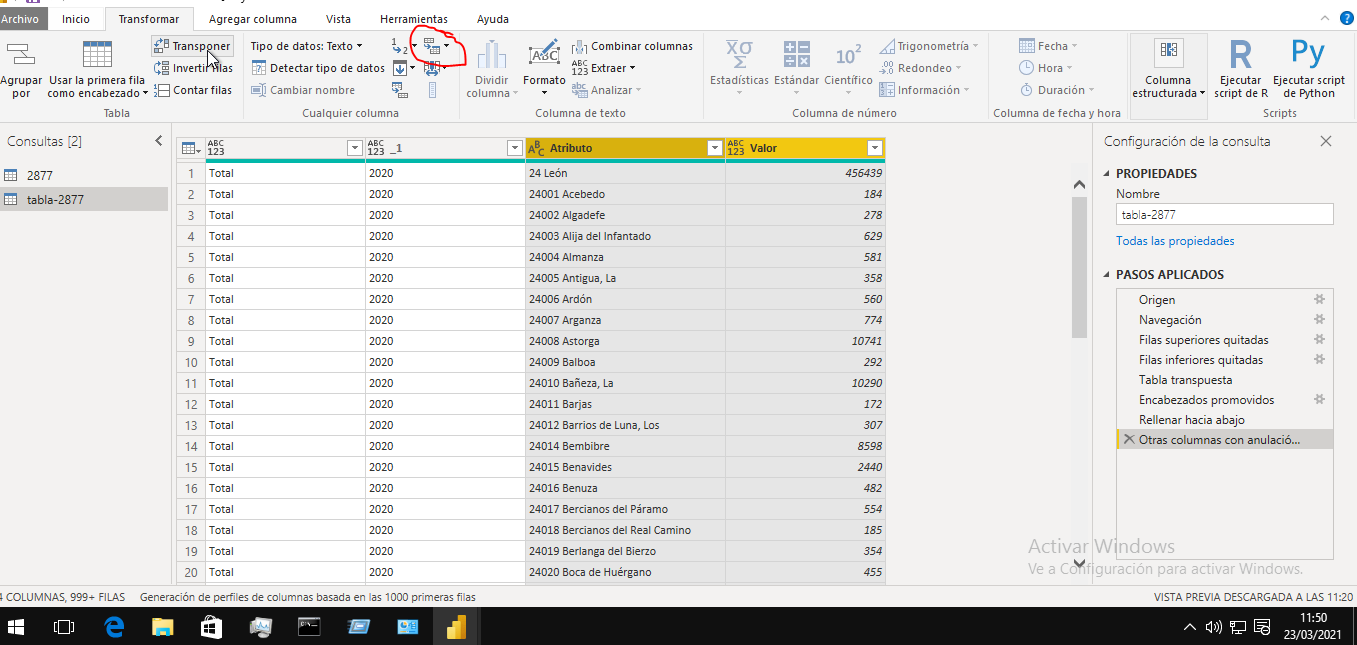


## Completar datos

Para que sea una tabla plana, debemos rellenar todos los datos que faltan(null). Para ello, damos click derecho sobre la cabecera de la columna y seleccionamos “Rellenar” y luego “abajo”.

## Eliminar dinamización de la tabla

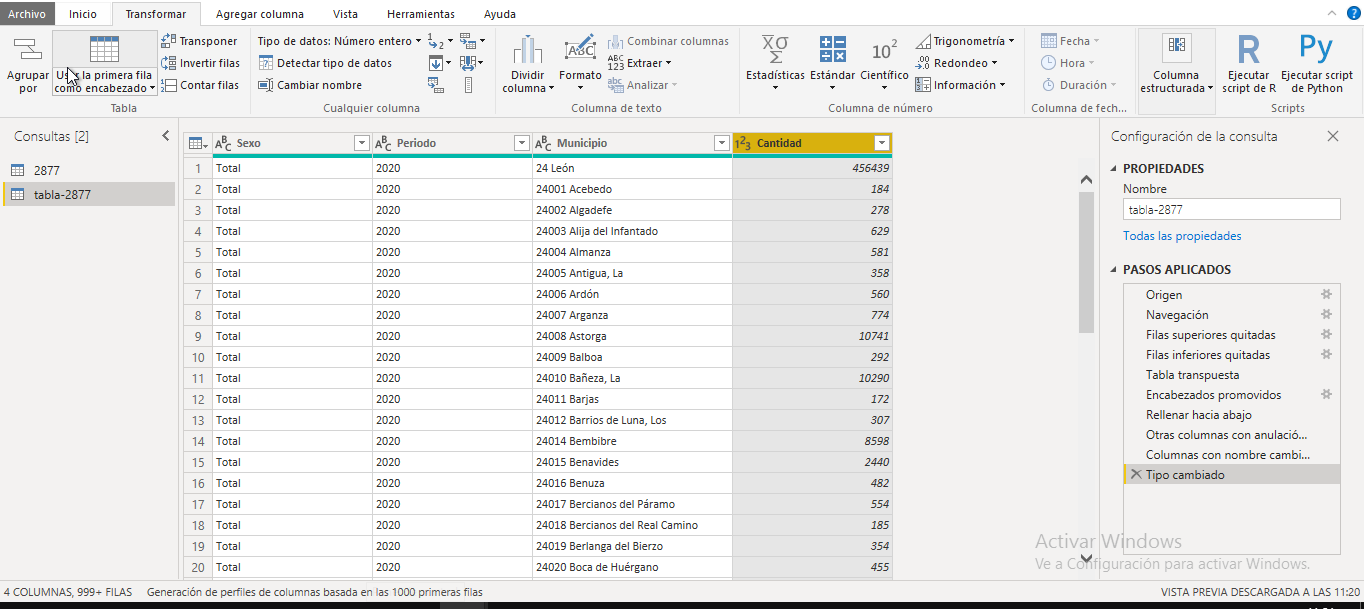
Pinchamos en las dos primeras(Control + click) -> Transformar -> “Anulación de dinamización de otras columnas”



## 1.5 Definir las columnas de la tabla

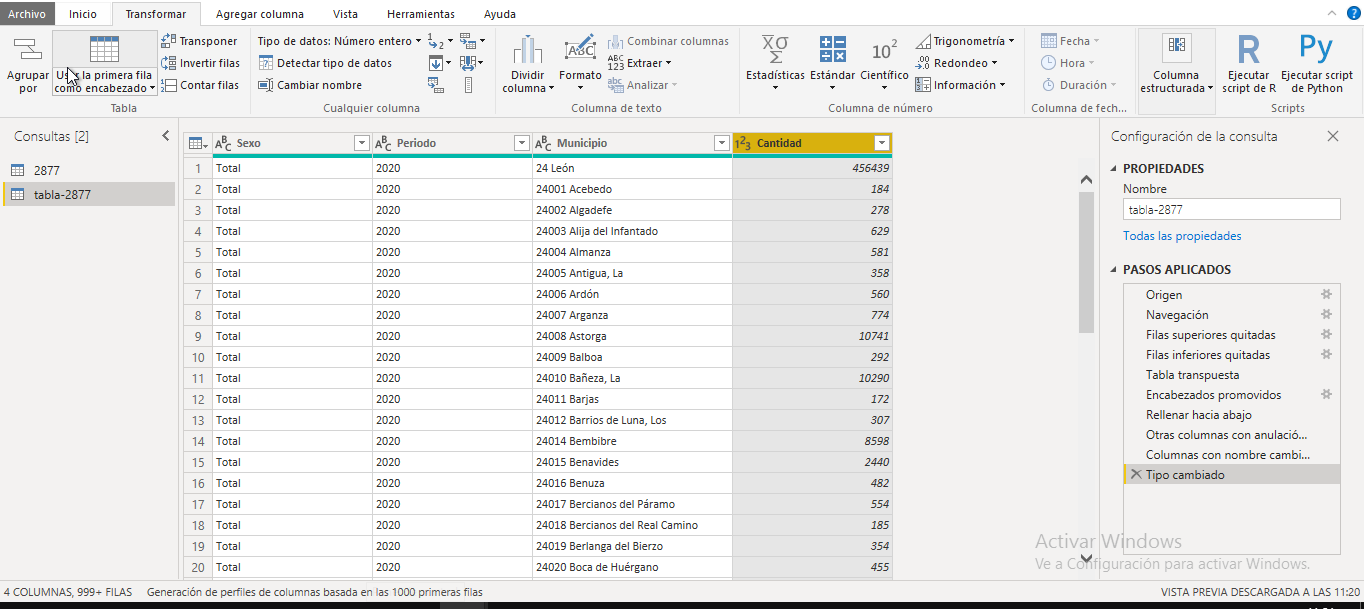
Ahora que tenemos los datos estructurados en forma de tabla plana, definimos los nombres de las columnas y sus tipos.

Todas serán de tipo “Texto” menos cantidad que será tipo “Número entero”.



### Ejercicio 3.1.

Importa y transforma tus datos en forma de tabla dinámica para obtener una tabla plana a partir de ellos (captura una pantalla del resultado final donde se vean todos los pasos del apartado PASOS APLICADOS y los nombres de los municipios).



## Definir los datos a nivel de individuo

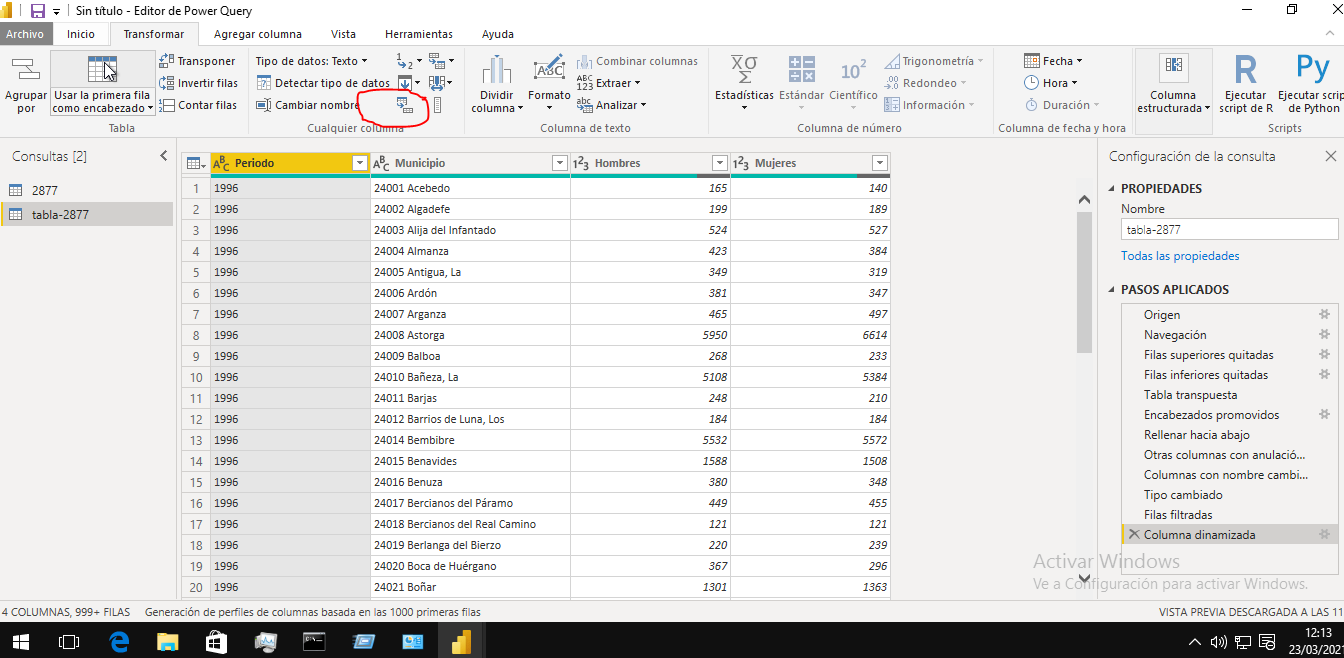
Pinchamos en la flechita de la columna “Municipios” y eliminamos la provincia(en mi caso León). Hacemos lo mismo en el caso de “Sexo”.

## Agrupar los datos a nivel de individuo

Como podemos ver, solamente tenemos una columna de sexo(hombre). Para diferenciar entre hombres y mujeres, la combinación se encuentra en Sexo y Cantidad.

Para ello, haremos una columna dinámica:

Transformar -> Cualquier columna -> Columna dinámica -> Seleccionamos Cantidad



## Obtener medidas adicionales de los individuos

Vamos a añadir el fichero “municipios.csv

1. Importar los datos

Nuevo origen -> CSV -> Elegimos el archivo

Eliminamos el paso aplicado “Tipo cambiado”

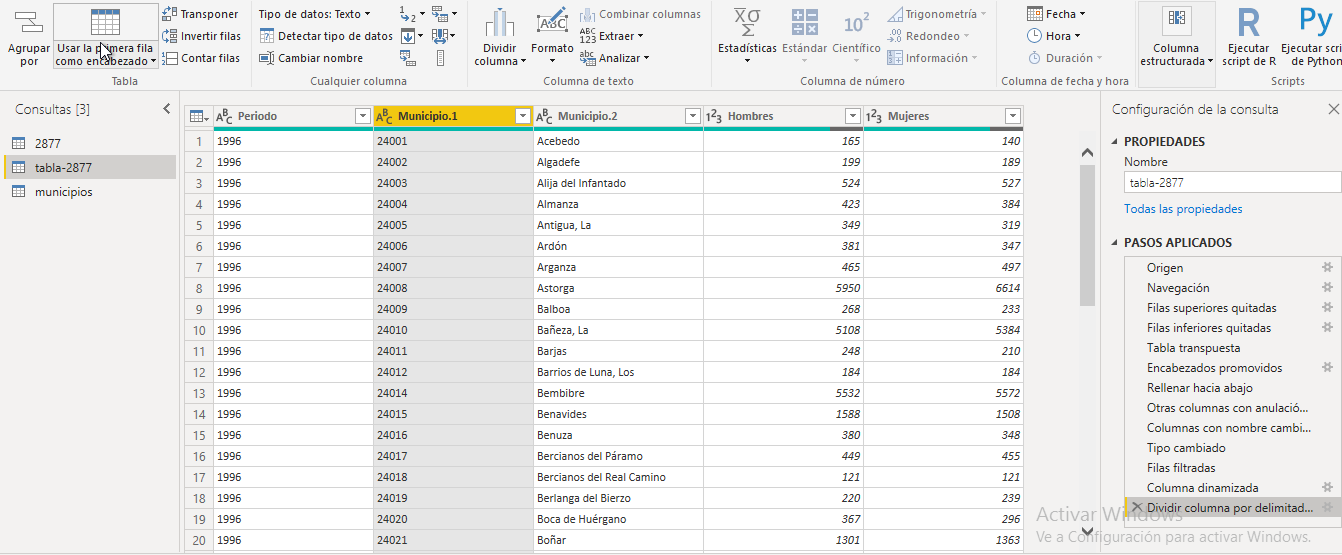
1. Adaptar los campos para tener uno en común

Juntamos los 5 primeros dígitos de “COD\_INE”. Para ello, separamos la columna Trasformar -> Columna de texto -> Dividir columna -> Por número de caracteres

Ponemos 5 y seleccionamos “Una vez, lo más a la izquierda posible”

También cambiamos el nombre de la nueva columna y eliminamos la terminación que se le ha incluido.

En la tabla original, también separamos el código postal del nombre del municipio(en este caso usamos “Delimitador” y seleccionamos Espacio(eliminamos Tipo cambiado)



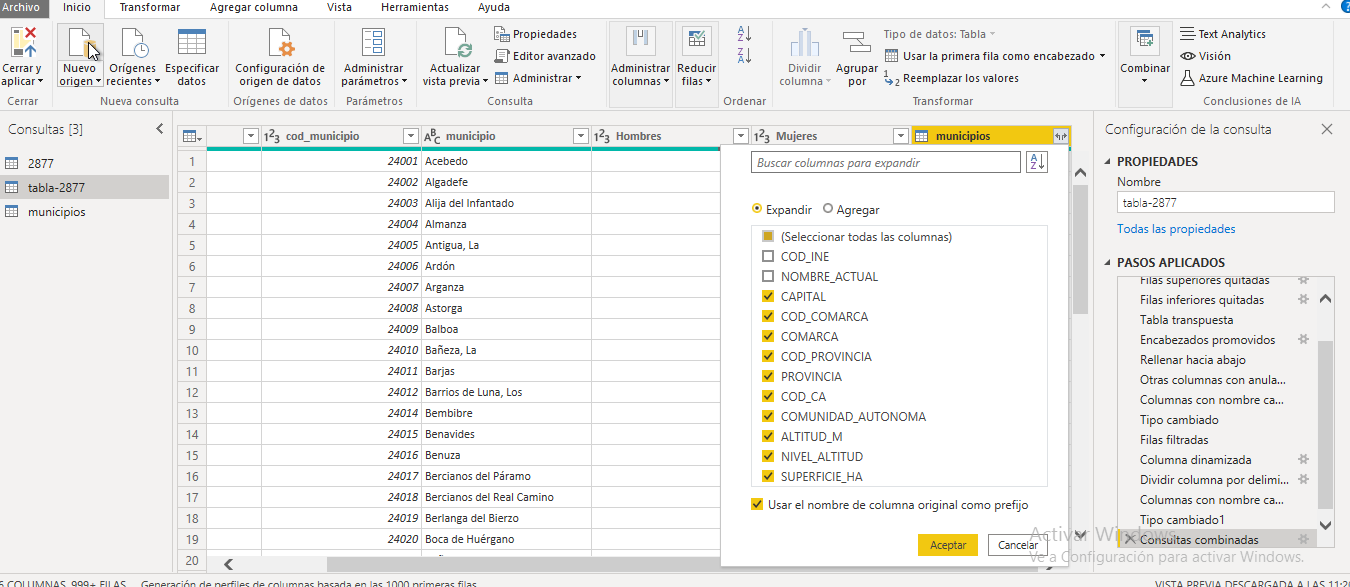
Renombramos los campos como cod\_municipio y municipio.

1. Definir la unión entre tablas

Agregamos los datos de la nueva tabla a la original.

Inicio -> Combinar -> Combinar consultas -> Combinas consultas

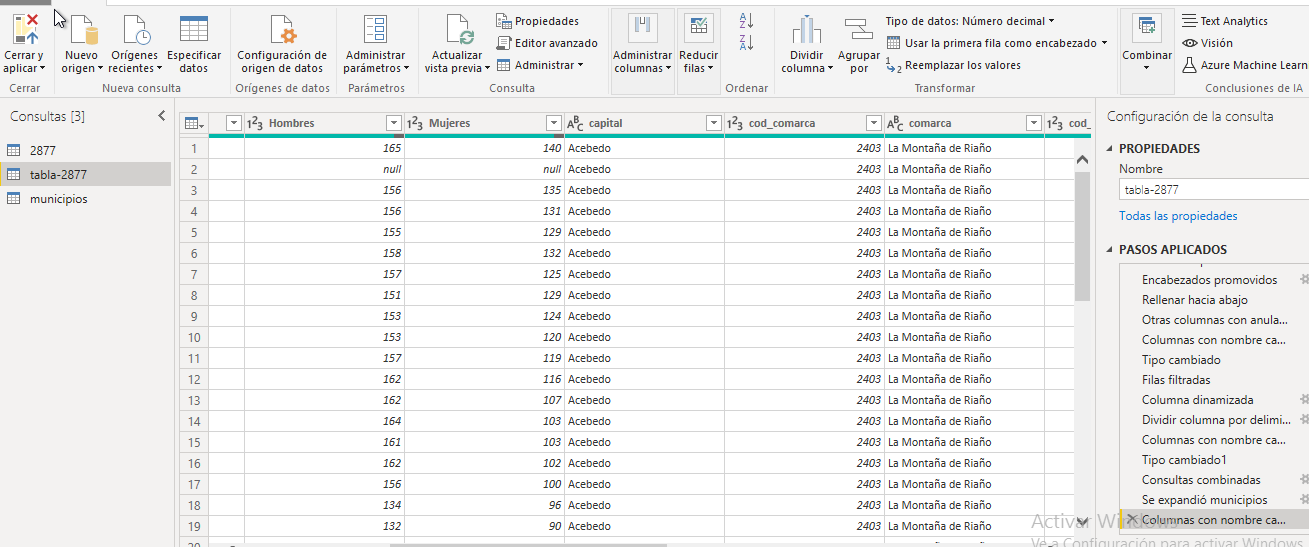
Seleccionamos cod\_municipio y en el desplegable seleccioamos COD\_INE. En tipo de combinación seleccionamos “Externa ixquierda”. Luego, en la nueva columna que se ha creado deseleccionamos COD\_INE y NOMBRE\_ACTUAL, como en la siguiente imagen. Estas dos columnas ya las tenemos.



Finalmente, cambiamos los nombres a minúsculas.

### Ejercicio 3.2

Define los datos de tu tabla a nivel de individuo y obtén mediciones adicionales de los individuos a partir de los datos del archivo municipios.csv, por último, selecciona y renombra los campos de la tabla resultante mediante el criterio snake case (captura una pantalla del resultado final donde se vean todos los pasos del apartado PASOS APLICADOS y los nombres de los municipios).

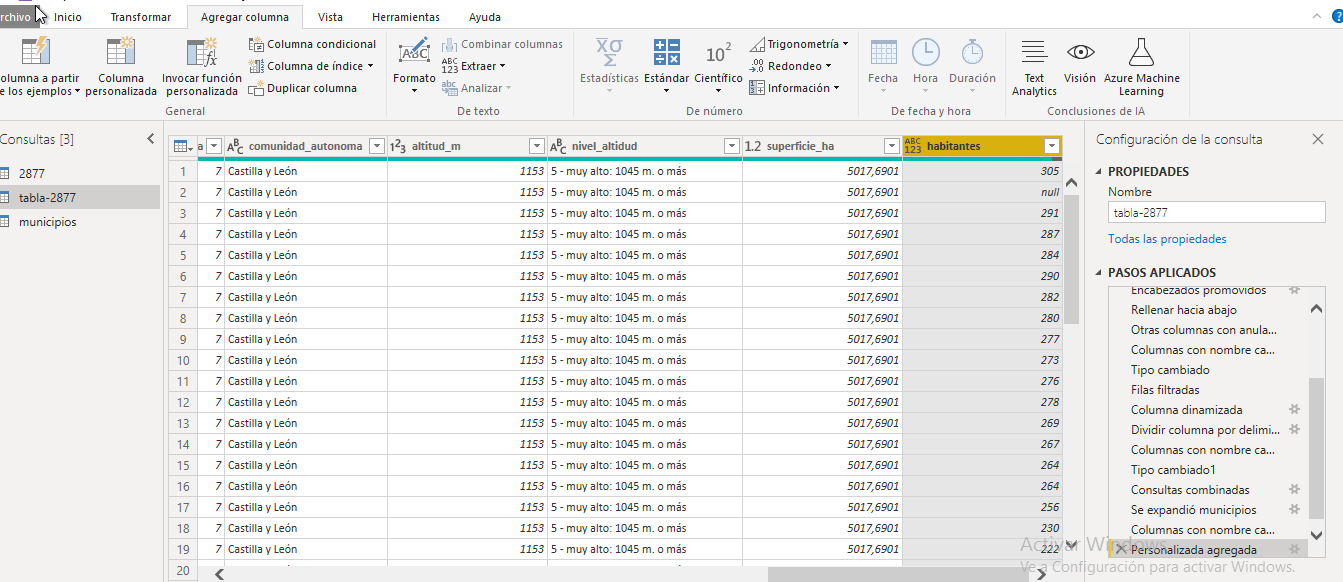


## Definir niveles adicionales de campos

Agregamos una nueva columna, que será la suma de las columnas “Hombre” y “Mujeres”.

Agregar columna -> General -> Columna personalizada

Ponemos como nombre “habitantes” y como fórmula [Mujeres] + [Hombres]

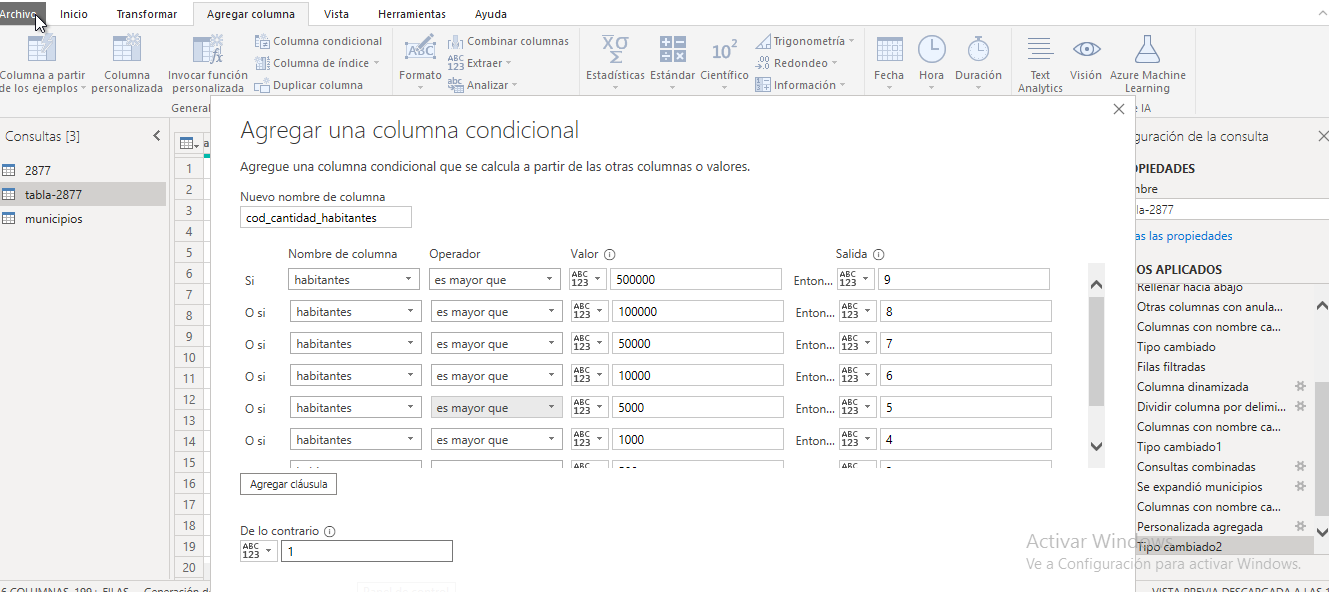


A la nueva columna le cambiamos el tipo a “Tipo entero”.

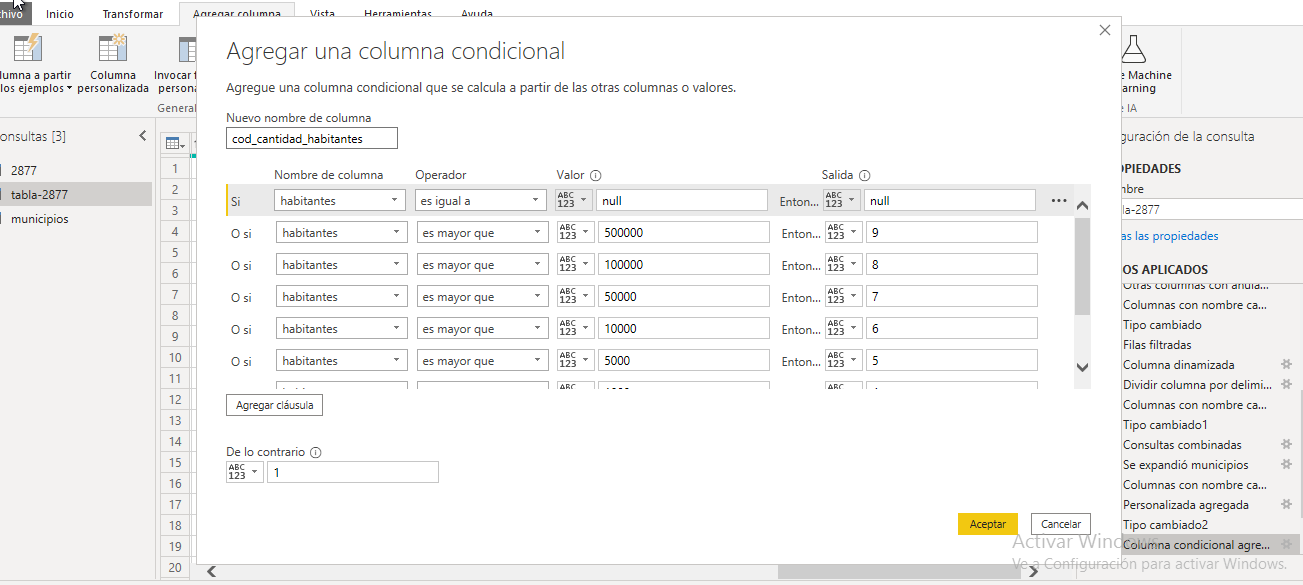
Ahora, agregaremos una columna adicional a partir de habitantes.

Nombre cod\_cantidad\_habitantes.

Los demás datos los iremos completando conforme el archivo

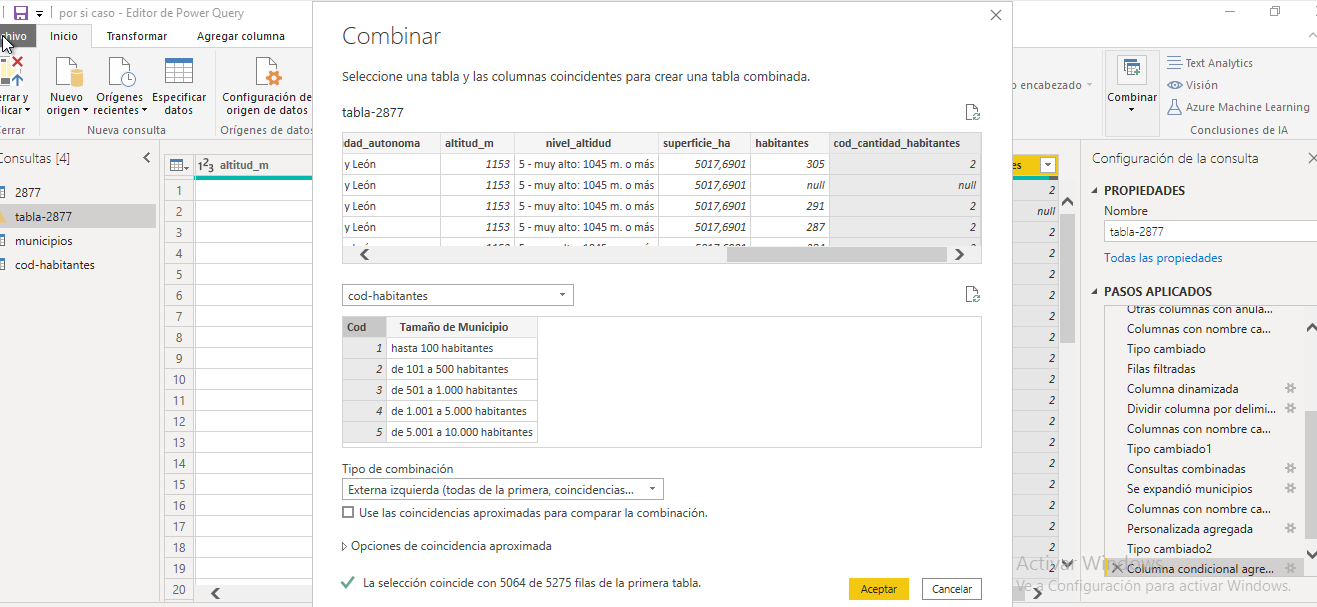


Para eliminar los “Error”, volvemos a editar este Paso aplicado y añadimos la siguiente clausula

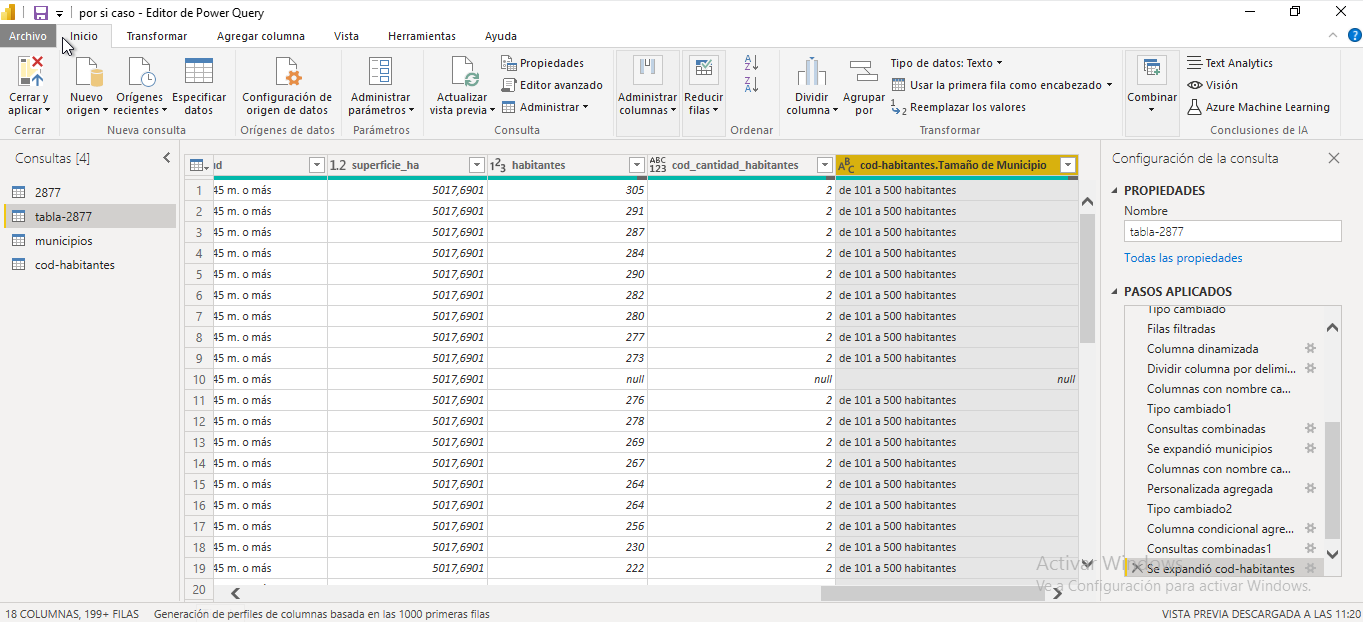


Muy importante que esté la primera.

El siguiente importar el archivo cod\_habitantes y combinarlo con la tabla de la siguiente forma:

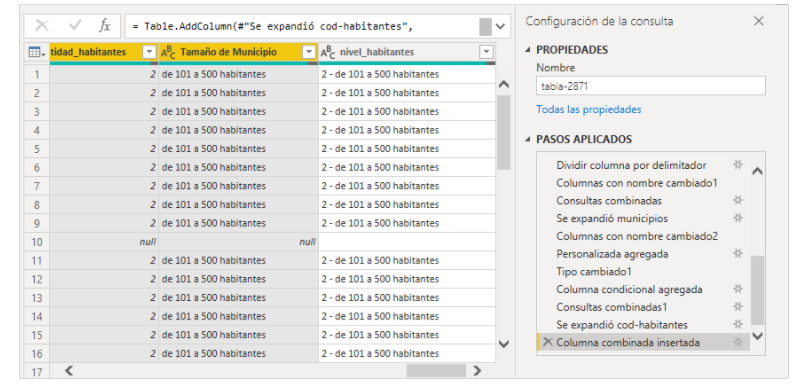


Seleccionando arriba la nueva columna(cod\_cantidad\_habitantes) y en la de abajo “Cod”. El resultado, tras deseleccionar el nivel Cod sería el siguiente:



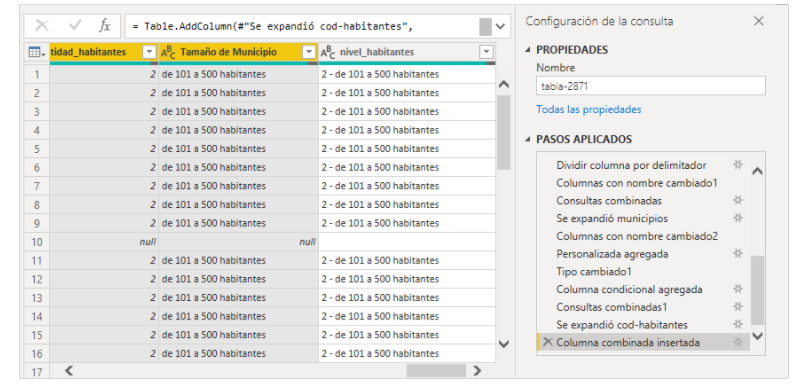
A continuación, combinamos código y descripción.

Agregar columnas -> De texto -> Combinar columnas con esta configuración



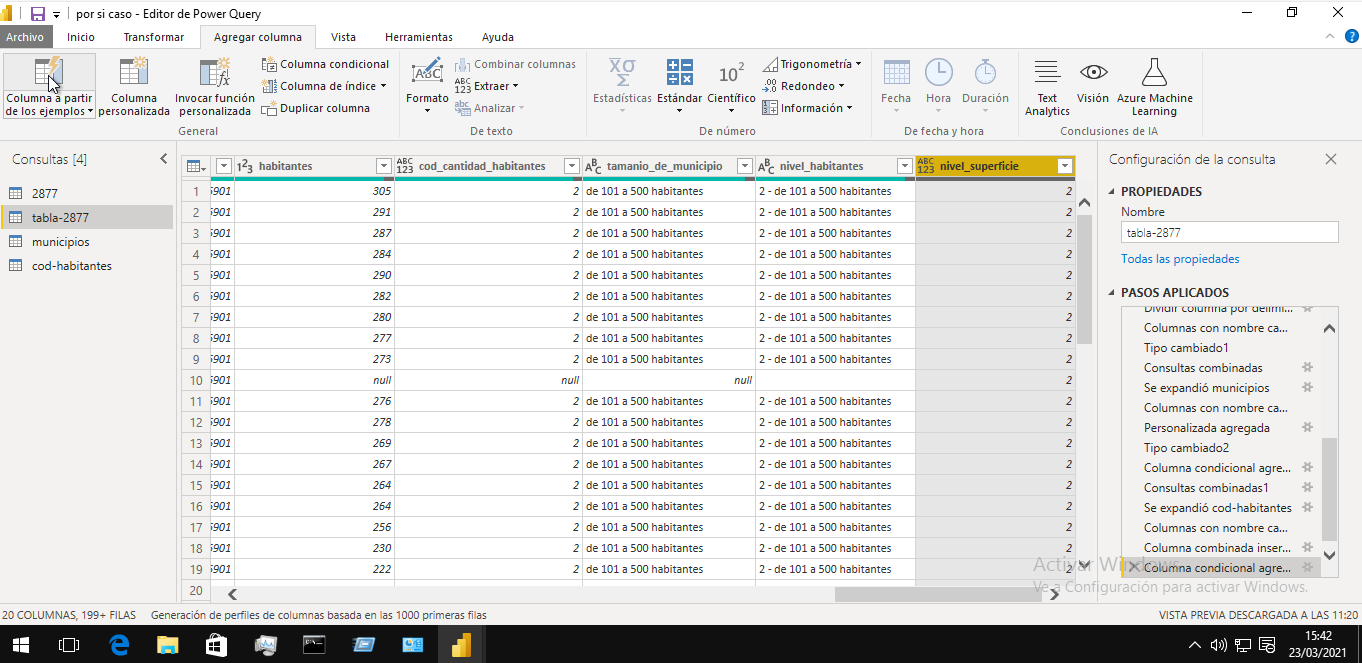
### Ejercicio 3.3

Define en la tabla original un campo llamado nivel\_habitantes con los códigos y descripciones del archivo cod-habitantes.csv definido en función de la cantidad de habitantes de cada municipio (captura una pantalla del resultado final donde se vean todos los pasos del apartado PASOS APLICADOS y los nombres de los municipios).



### Ejercicio 3.4

Define en la tabla original un campo llamado nivel\_superficie en función del campo superficie\_ha, similar a los anteriores[13](https://lsi2.ugr.es/jsamos/sm2019/power-query.html#fn13), con 3 niveles, con los límites que consideres oportunos según tu criterio (captura una pantalla del resultado final donde se vean todos los pasos del apartado PASOS APLICADOS y los nombres de los municipios, y otra donde se vea la definición de los niveles).



### Generar hechos y dimensiones

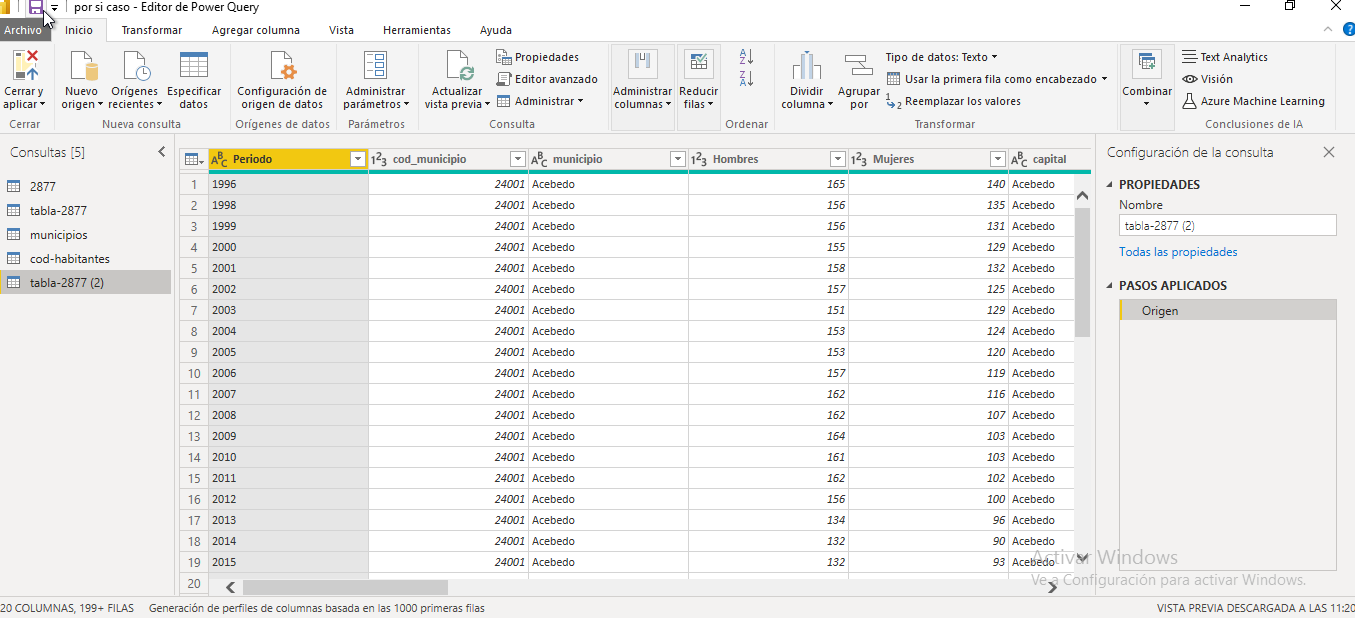
Eliminamos los valores null de todas las columnas, yendo una a una comprobando.

## Generar las dimensiones

Generaremos consultas mediante referencia.

Dimensión Cuándo

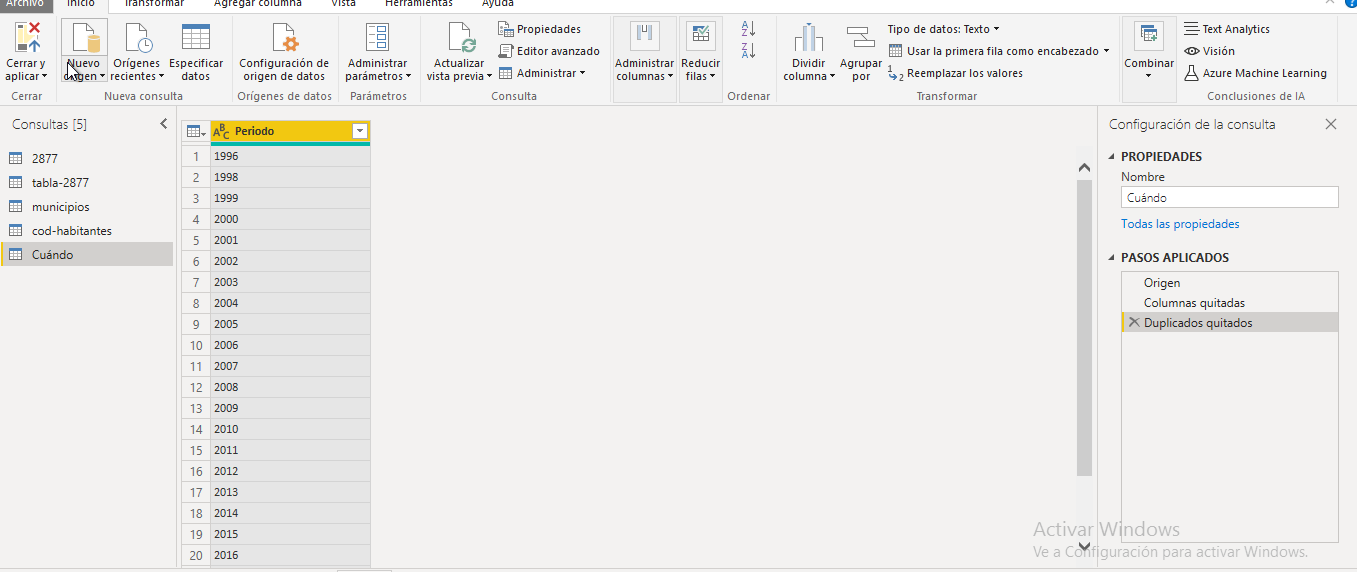
Inicio -> Consulta -> Administrar -> Referencia



1. ´Renombramos y le llamamos “Cuándo” y eliminamos todas las columnas menos el periodo.

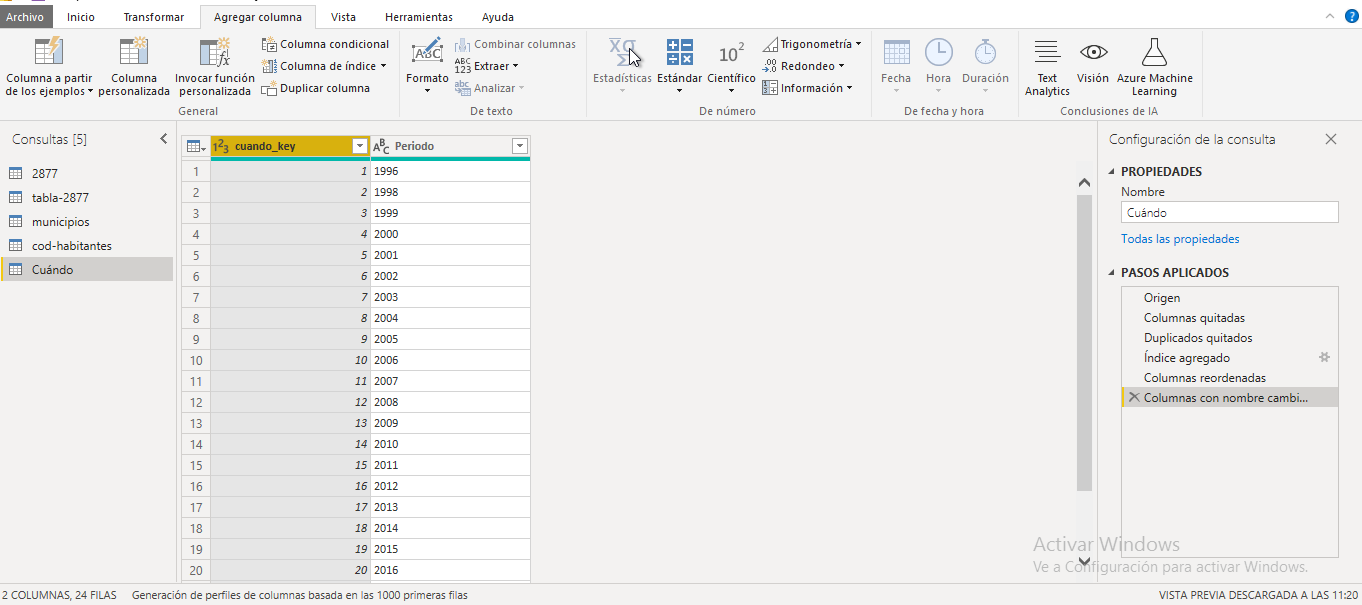
Inicio -> Administrar columnas -> Quitar columnas -> Quitar otras columnas

Seleccionamos todas menos el periodo y pulsamos click. También eliminaremos duplicados.



1. Añadimos una clave generada.

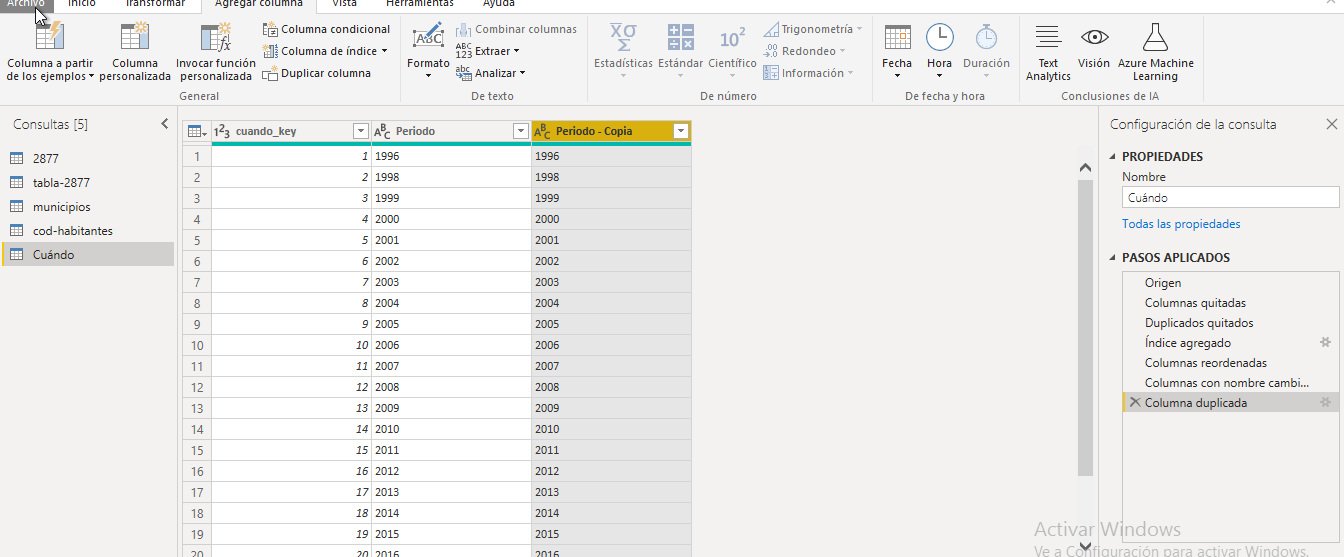
Agregar columna -> General -> Columna de índice



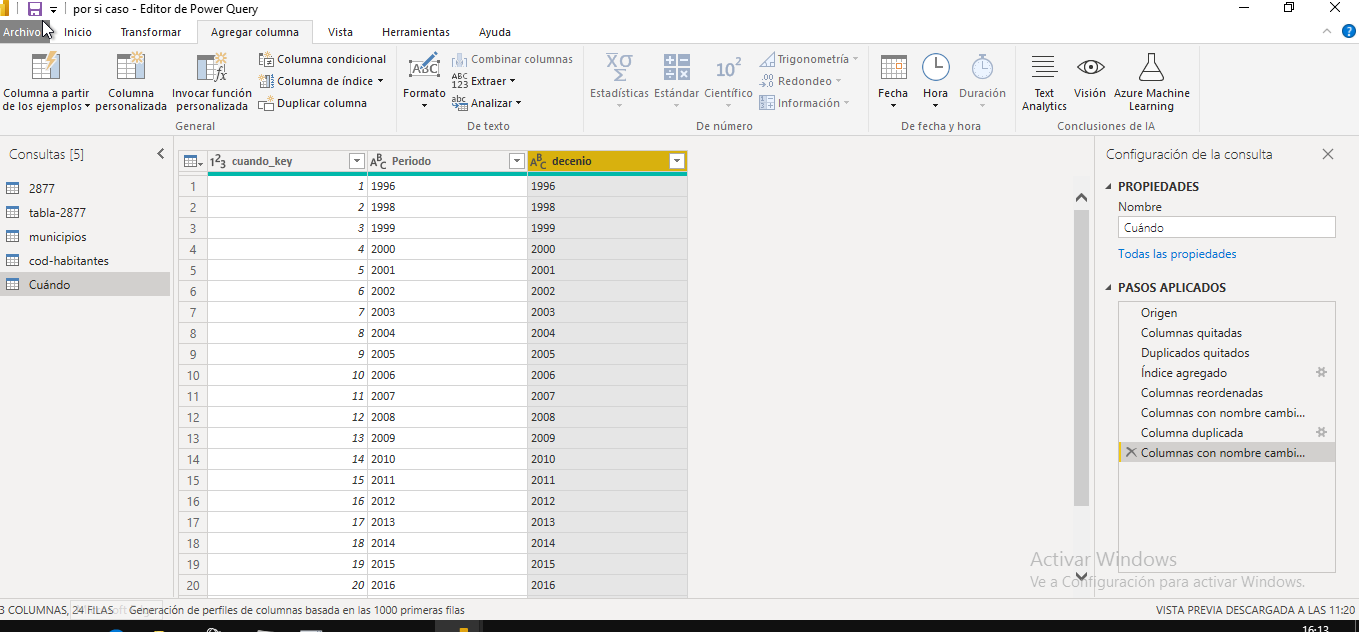
1. Añadimos un campo “decenio”.

Seleccionamos “Periodo” y pulsamos Agregar columna -> Columna a partir de los ejemplos -> A partir de la selección

Iniciamos a copiar los datos y nos autocompletará los campos.

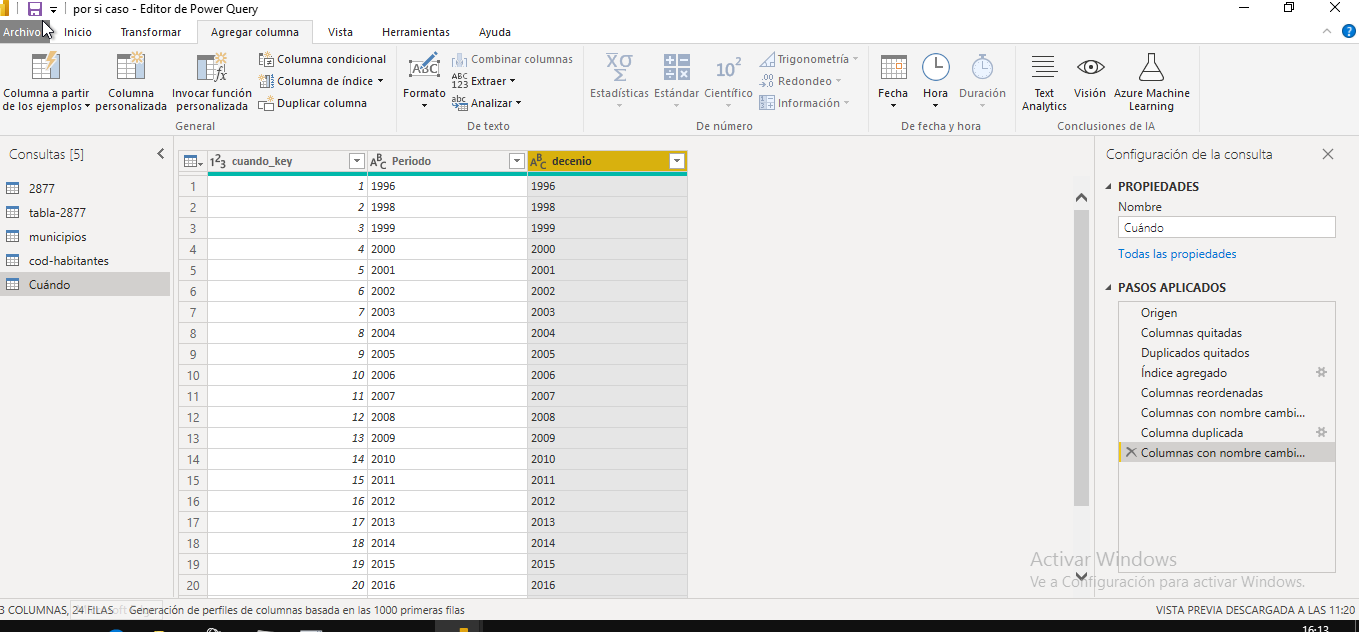


1. Cambiamos el nombre del campo por “decenio”



### Ejercicio 3.5

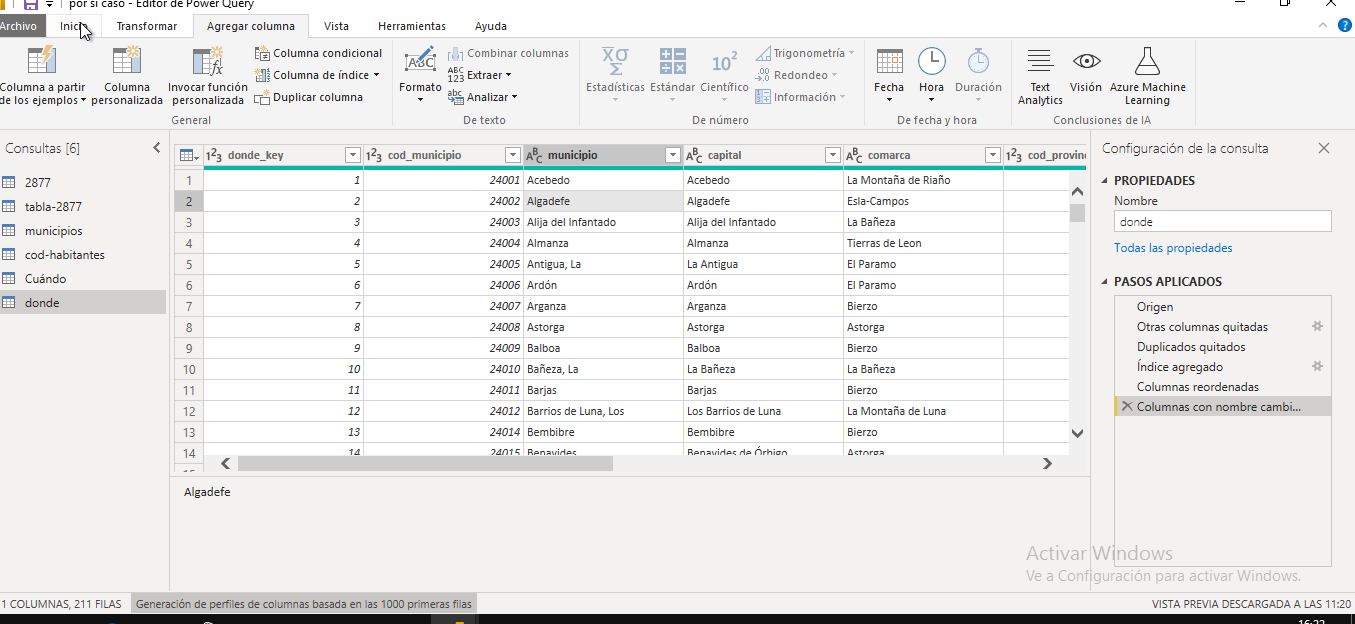
Genera la tabla de la dimensión Cuándo con los campos periodo, decenio y una llave generada (captura una pantalla del resultado final donde se vean todos los pasos del apartado PASOS APLICADOS).



Dimensión Dónde

Hacemos lo mismo pero con todos los campos menos periodo,hombres,mujeres,cod\_comarca,cod\_comunidad\_autonoma,altidud\_m,etc.

Quitamos duplicados, agregamos índice, la colocamos al principio y le cambiamos el nombre.



### Ejercicio 3.6

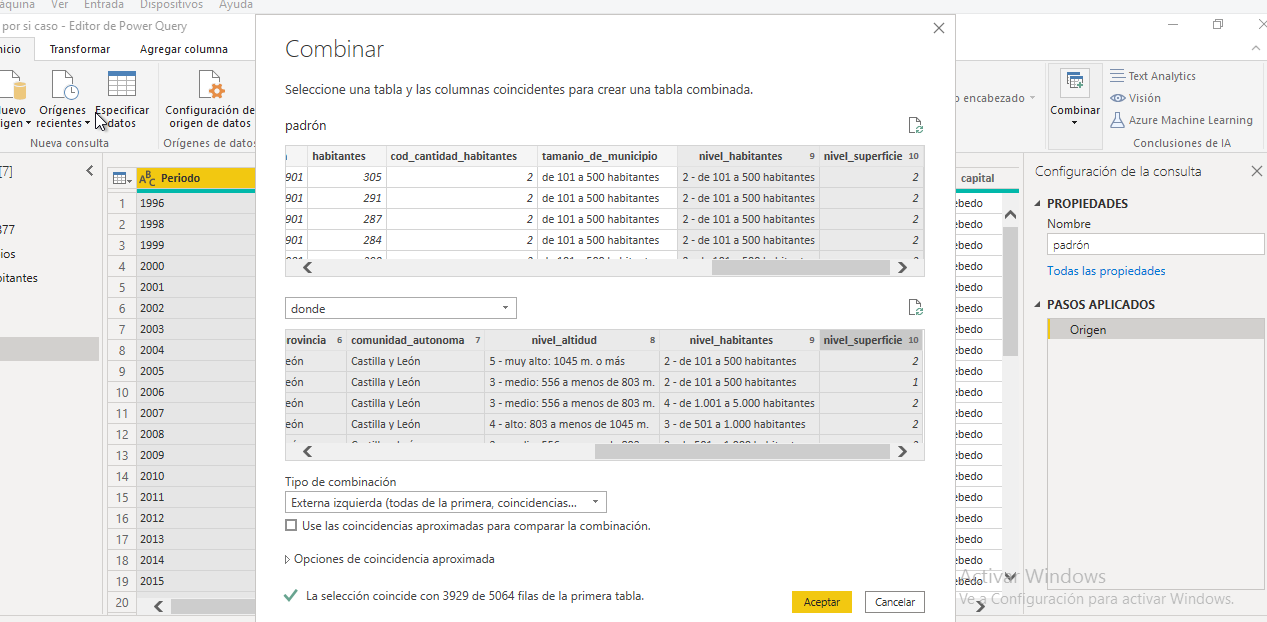
Genera la tabla de la dimensión Dónde con los campos propios de esa dimensión (al menos los incluidos en este apartado) y una llave generada (captura una pantalla del resultado final donde se vean todos los pasos del apartado PASOS APLICADOS).

* En mi caso, no existen duplicados.
* Sucedería que habría dos claves iguales y, o da error, o devuelve valores erróneos.

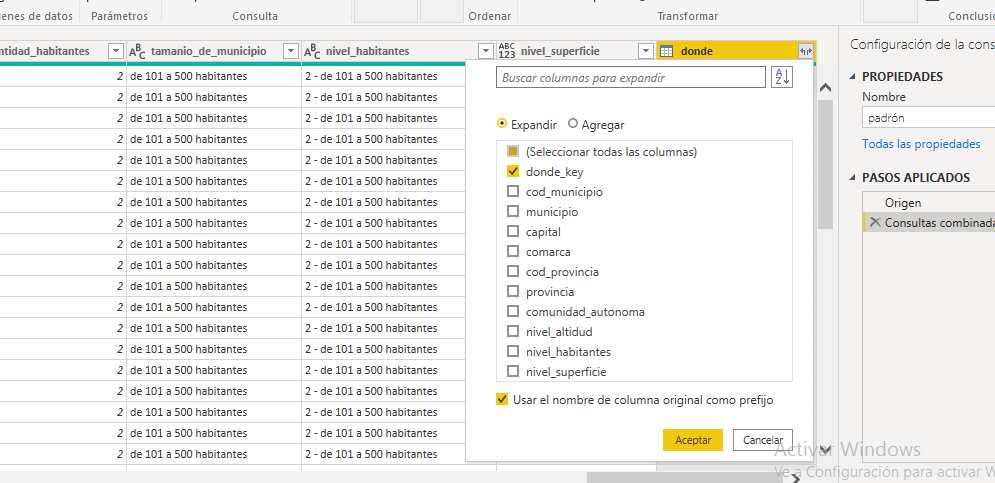
## 3.5.2 Generar los hechos

Creamos una nueva consulta, que llamaremos “padrón”. Para ello, pinchamos en Inicio -> Combinar consultas -> Combinar consultas

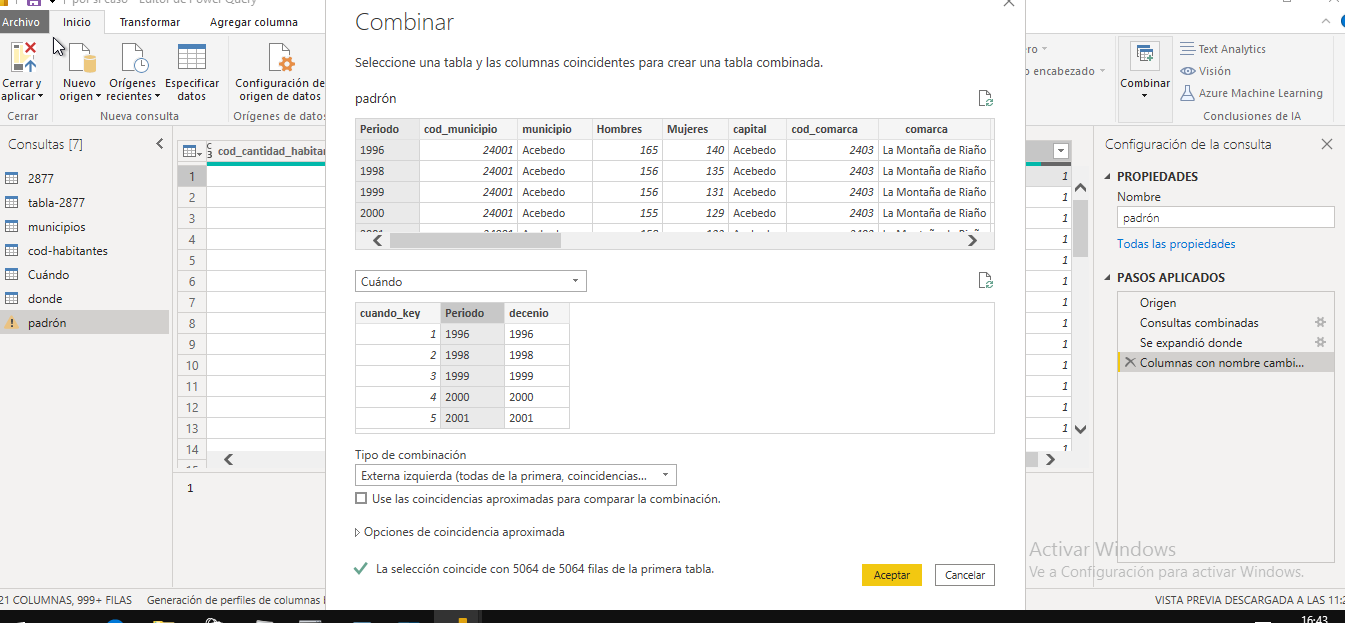
En la siguiente pantalla, empezamos por la consulta dónde, seleccionando en ella las que estén en la principal EN ORDEN.



Comprobaremos si se ha establecido correctamente la relación. Al expandir, seleccionaremos únicamente la llave generada donde\_key

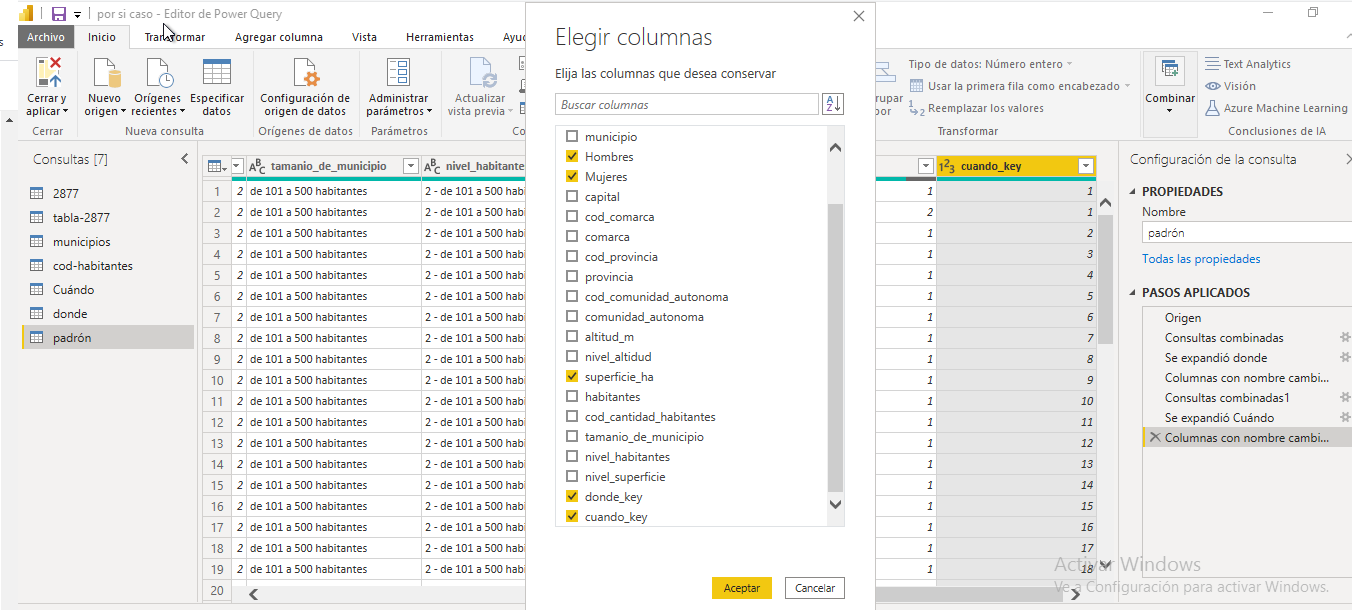


Hacemos lo mismo con la Cuando

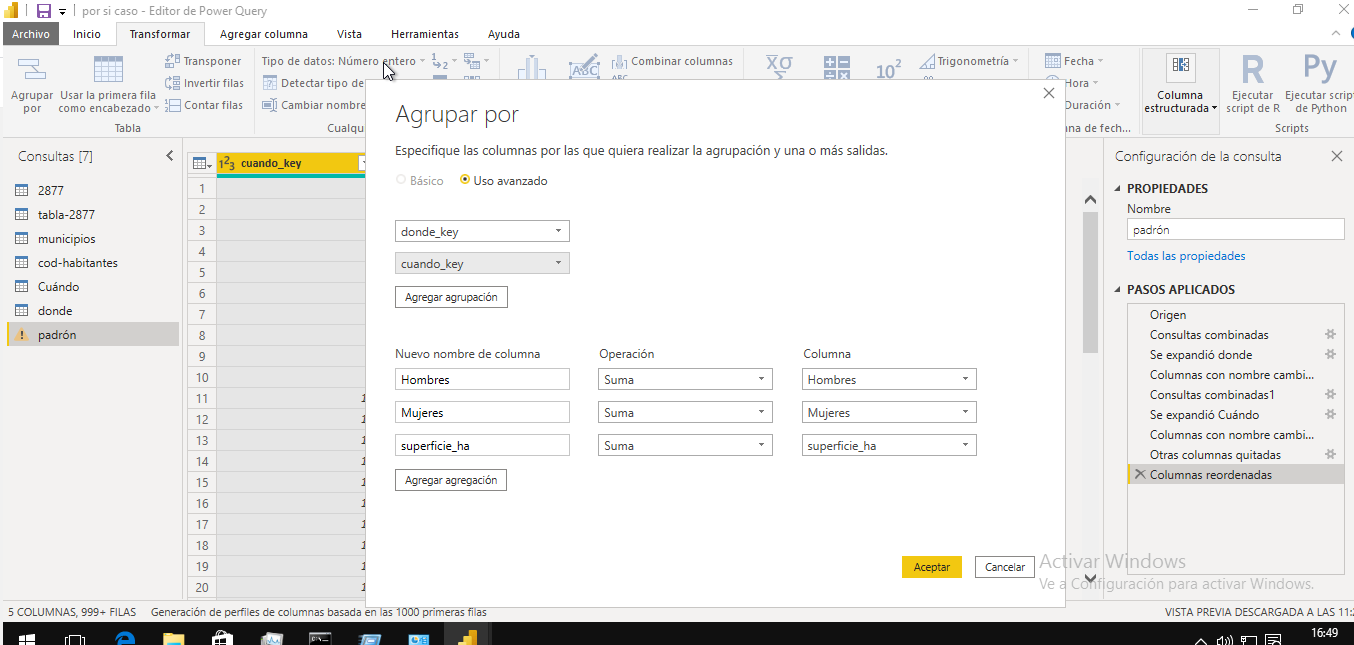


Para terminar, debemos seleccionar las llaves junto a la medidas.

Inicio -> Administrar columnas -> Elegir columnas -> Elegir columnas



Agrupamos por si se ha fragmentado algo.



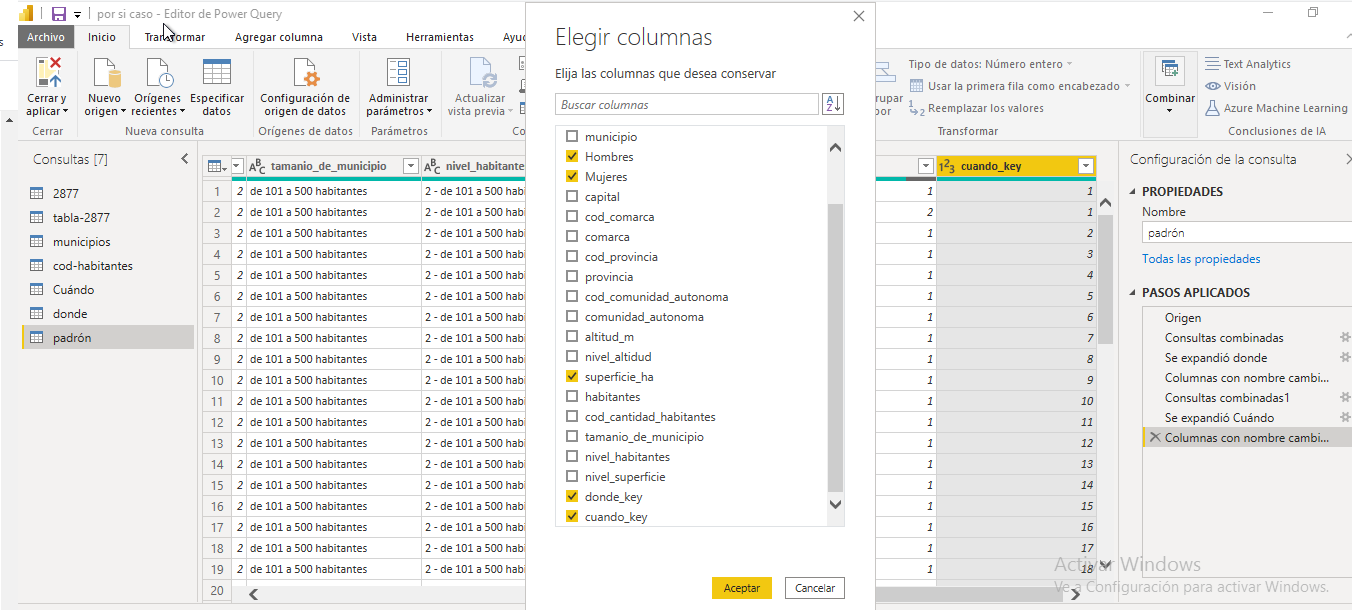
Si hacemos esto, deberemos cambiar el tipo, ya que se habrá cambiado a decimal.

## Seleccionar las tablas que se devuelven a Power BI

Hacemos click derecho en tabla-2971, municipios y cod-habitantes y pulsamos en habilitar carga, para no enviarlas a power BI.

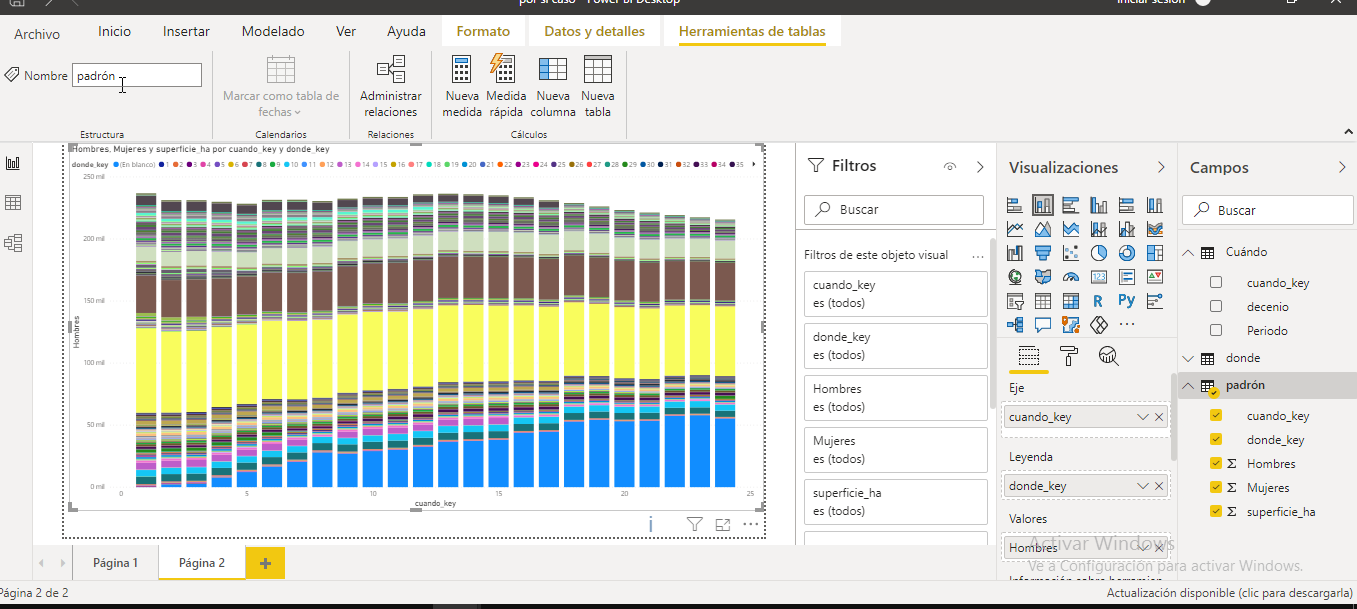
### Ejercicio 3.7

Genera la tabla de hechos padrón (captura una pantalla del resultado final donde se vean todos los pasos del apartado PASOS APLICADOS).



Si faltan datos, anteriormente explico paso a paso cómo lo he hecho.

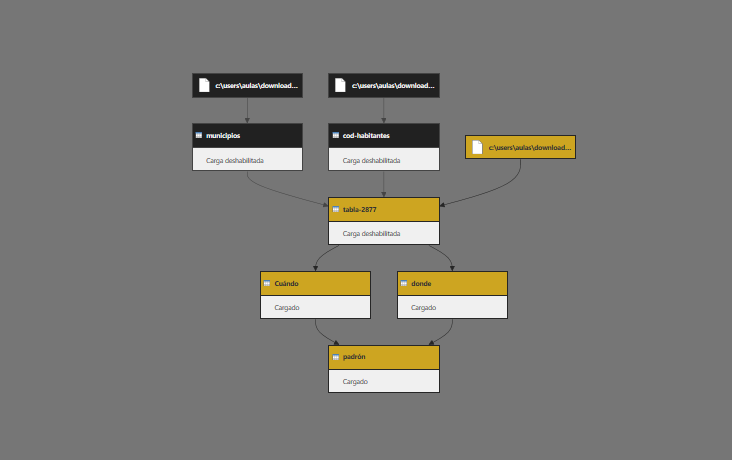
* En la definición de la unión entre la tabla con todos los datos y la tabla de la dimensión Dónde se han usado todos los campos en común. ¿Cuál sería el conjunto mínimo de campos a utilizar en la unión para obtener el mismo resultado y por qué?
* Configura las consultas definidas de manera que solo pasen a Power BI las tablas de hechos y dimensiones, y actualízalas en Power BI (captura una pantalla de Power BI donde se vean los datos disponibles para consulta).



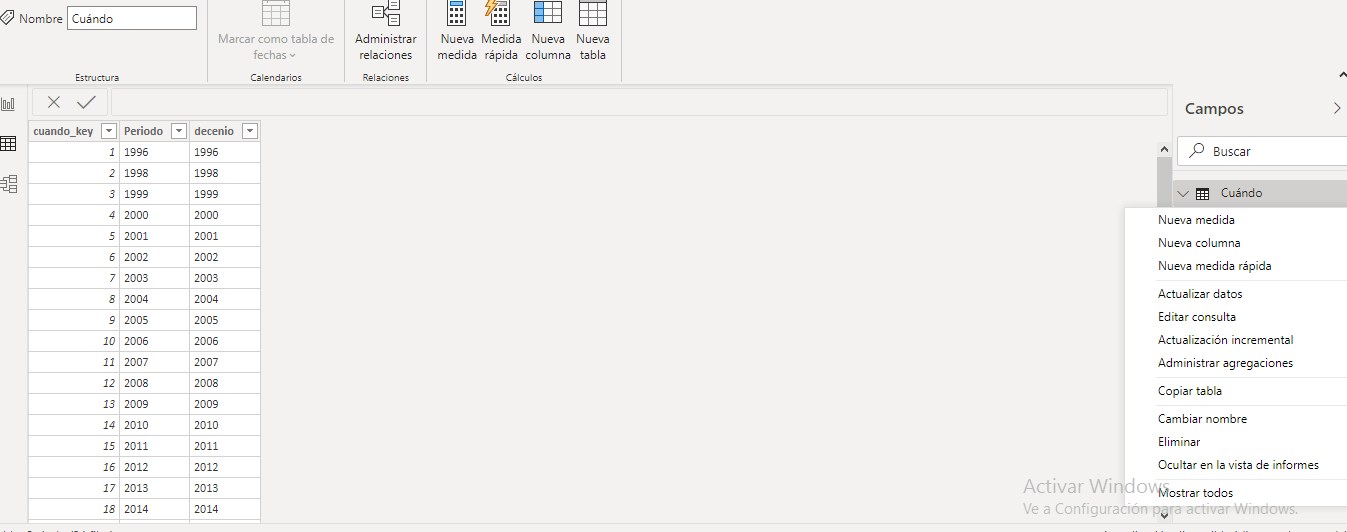
### Ejercicio 3.8

Obtén la siguiente información adicional:

* Representación gráfica de las transformaciones realizadas (captura la pantalla).



* Guarda las tablas de dimensiones y hechos en formato CSV (captura una pantalla para cada tabla en la hoja de cálculo). **Conserva estos archivos por si fuese necesario usarlos en prácticas posteriores.**

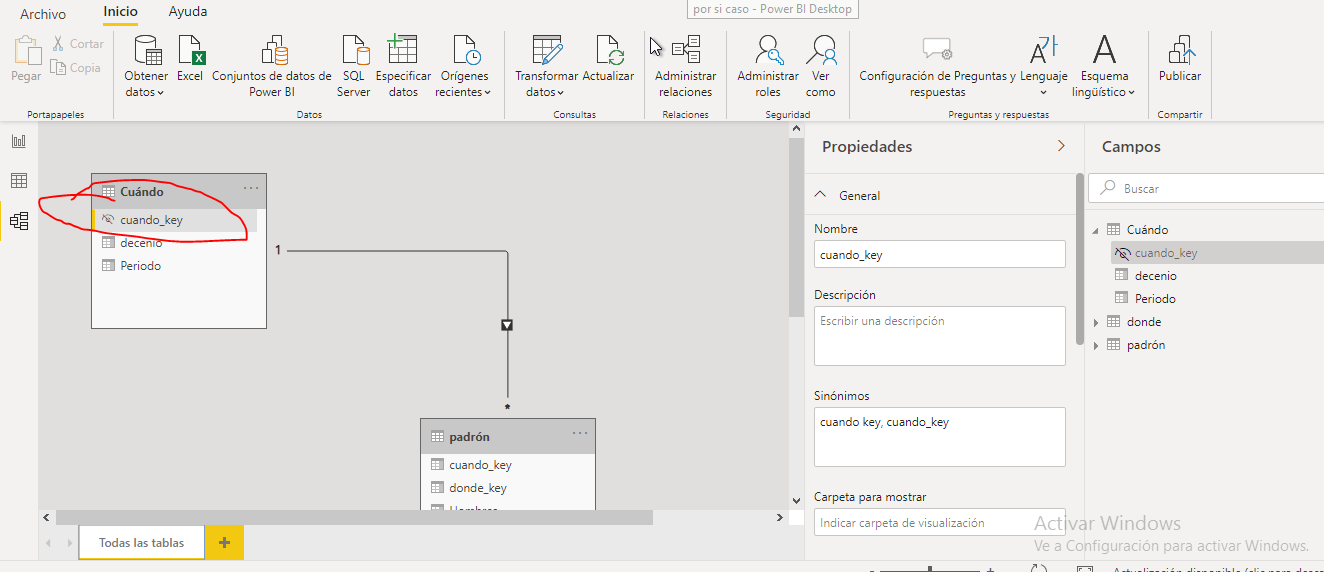


Me es imposible copiar el archivo a Excel, ya que en esta imagen se incluye. Uso virtual box para lanzarla y partiendo de que va extremadamente lenta, me resultaría casi imposible instalar el paquete office.

Hago captura enseñando que sé cómo se hace.

# Modelo y consultas en Power BI

Primero, ocultaremos las llaves generadas en la vista de los informes, ya que no son necesarios para ver las relaciones entre las tablas.



Botón derecho encima y “Ocultar en la vista de informes”

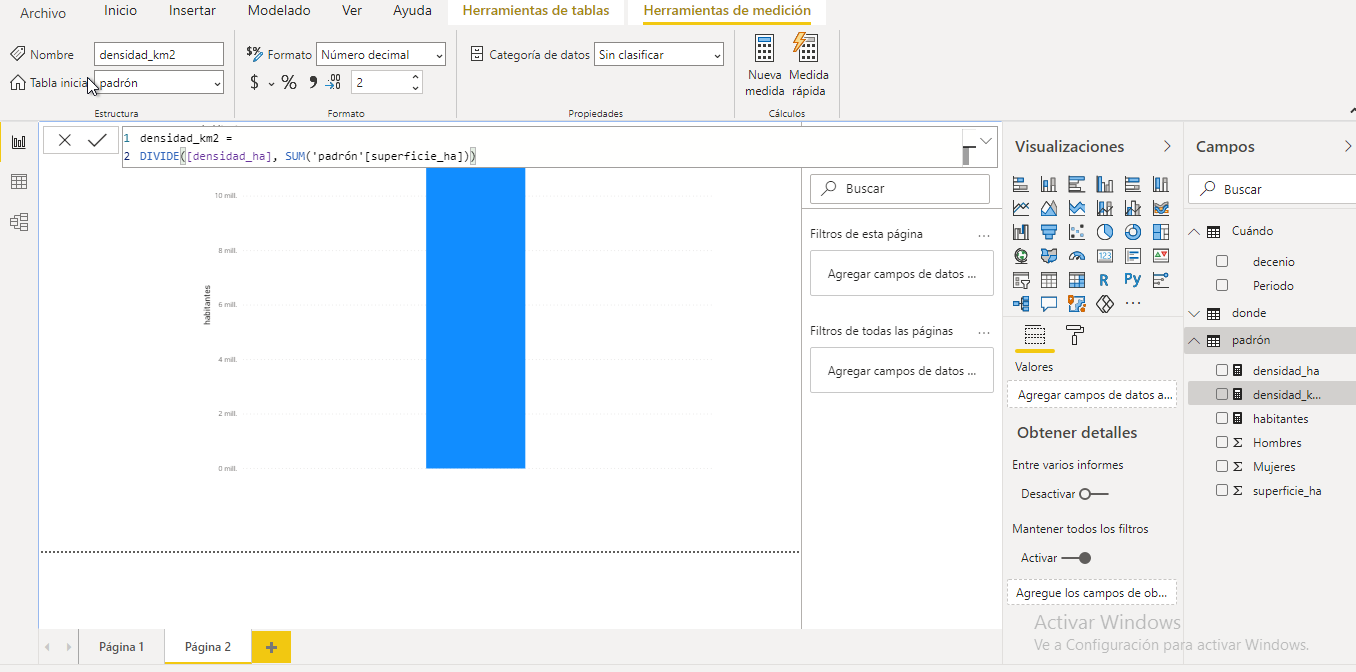
## Definición de medidas calculadas

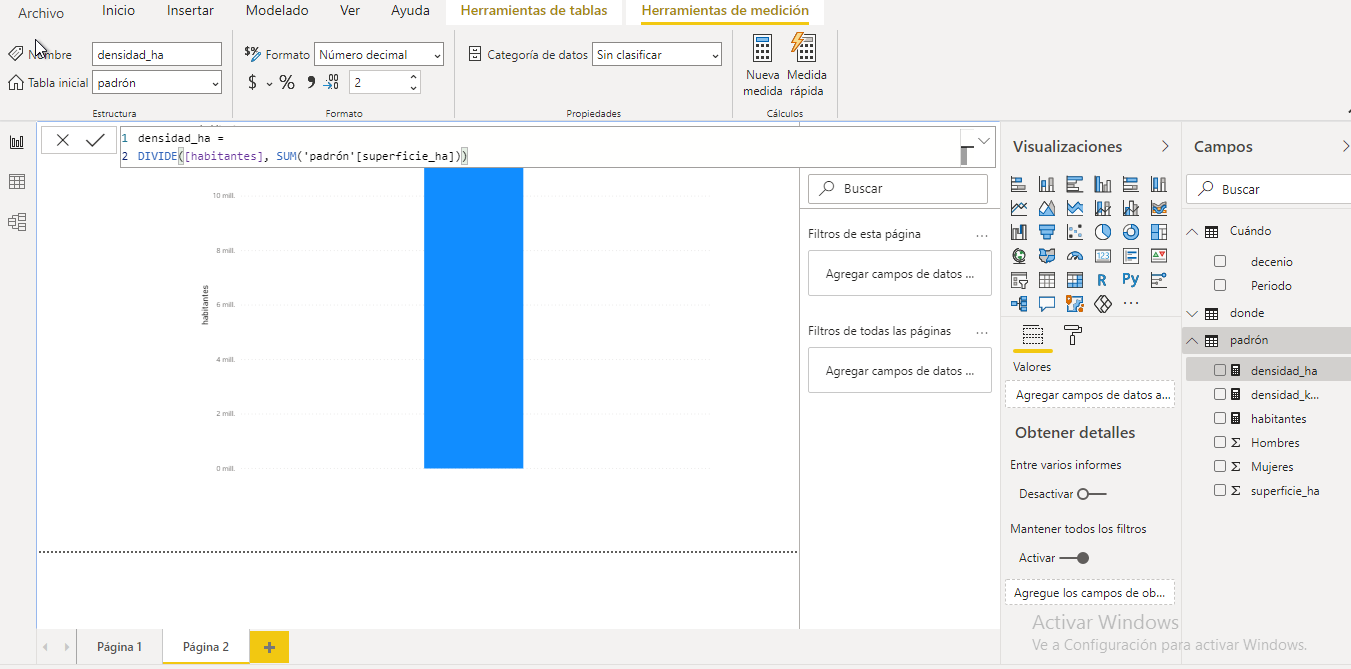
Vamos a definir las medidas calculadas mediante DAX.

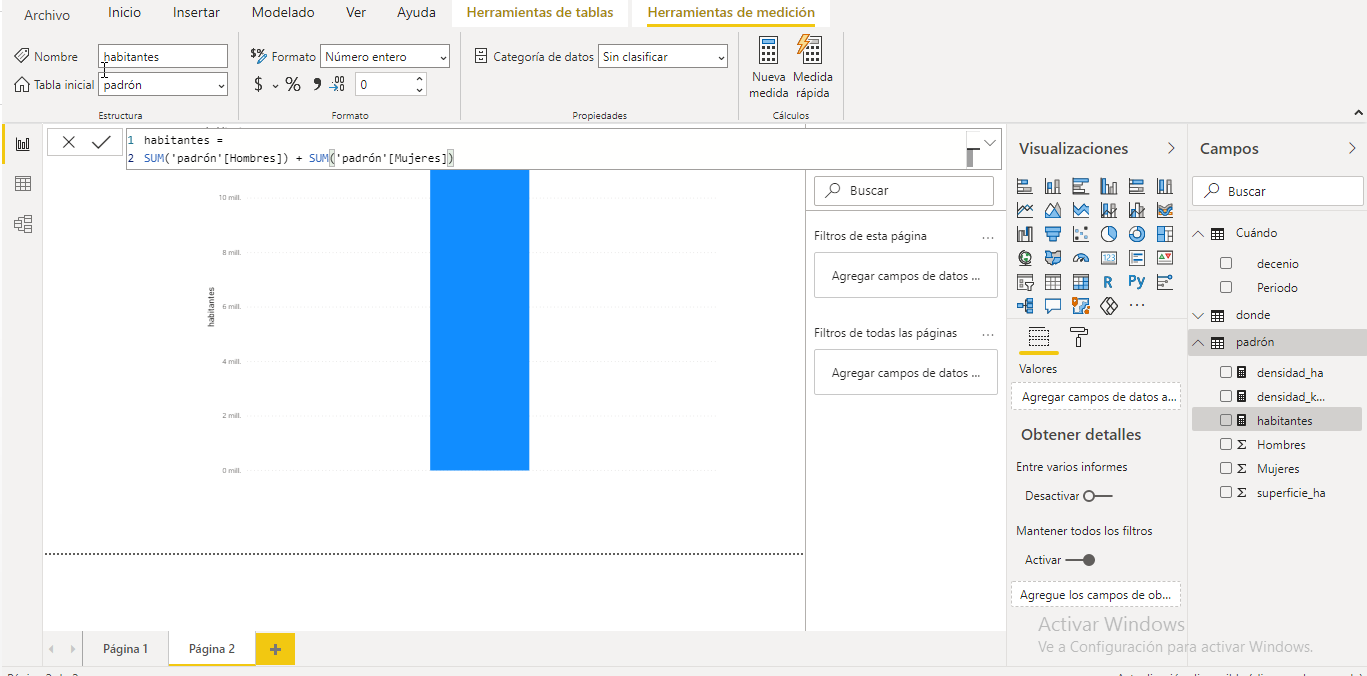
Para definir una vista vamos a la vista “Informe” y pulsamos Inicio -> Cálculos -> Medida rápida

### Ejercicio 4.1

Oculta las llaves generadas y define las medidas calculadas habitantes, densidad\_ha y densidad\_km2 según las explicaciones previas (captura una pantalla para cada medida donde se muestre su definición con la fórmula en DAX).

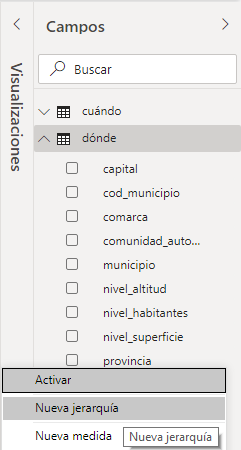




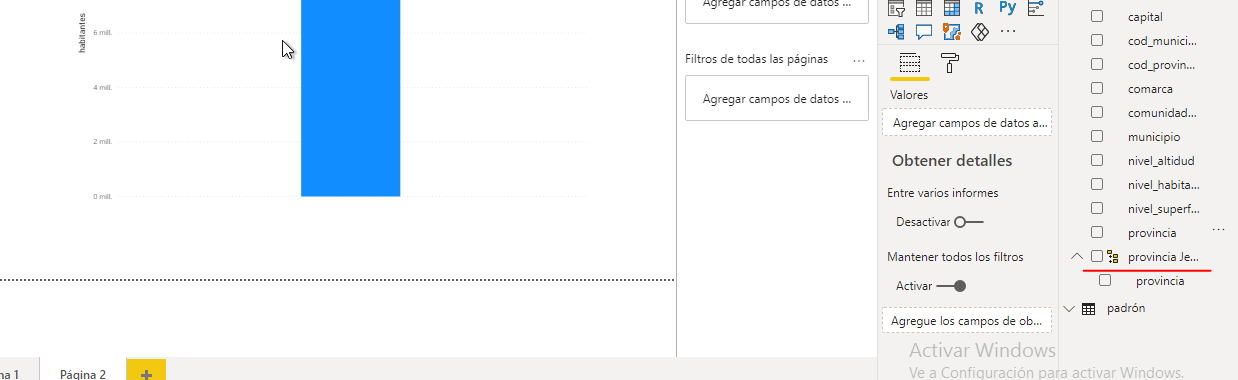


## Definición de jerarquías

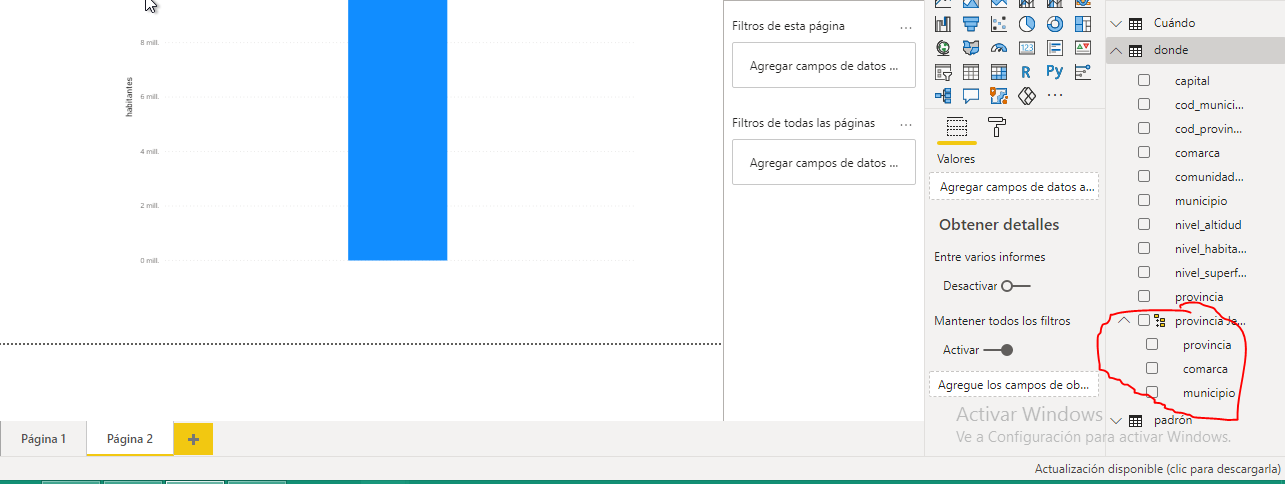
Definimos una jerarquía desde provincia hasta municipio pasando por comarca. Lo hacemos de esta forma



Al hacer esto, aparece un nuevo campo

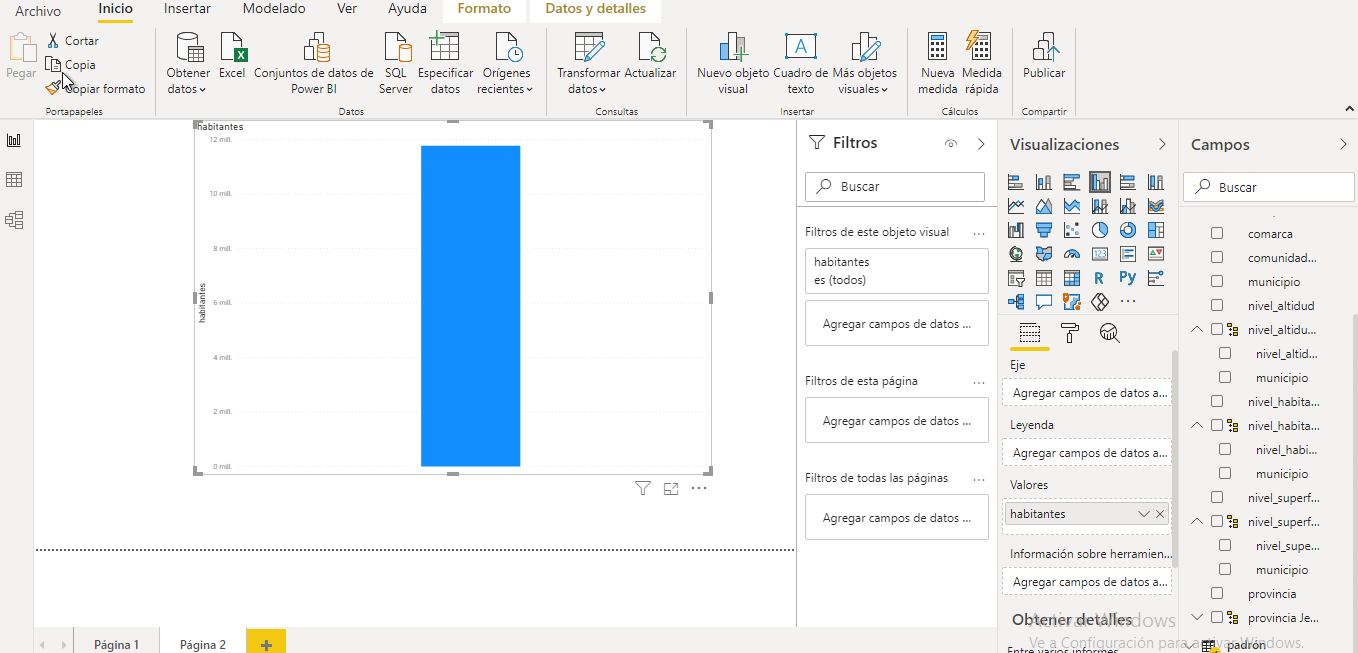


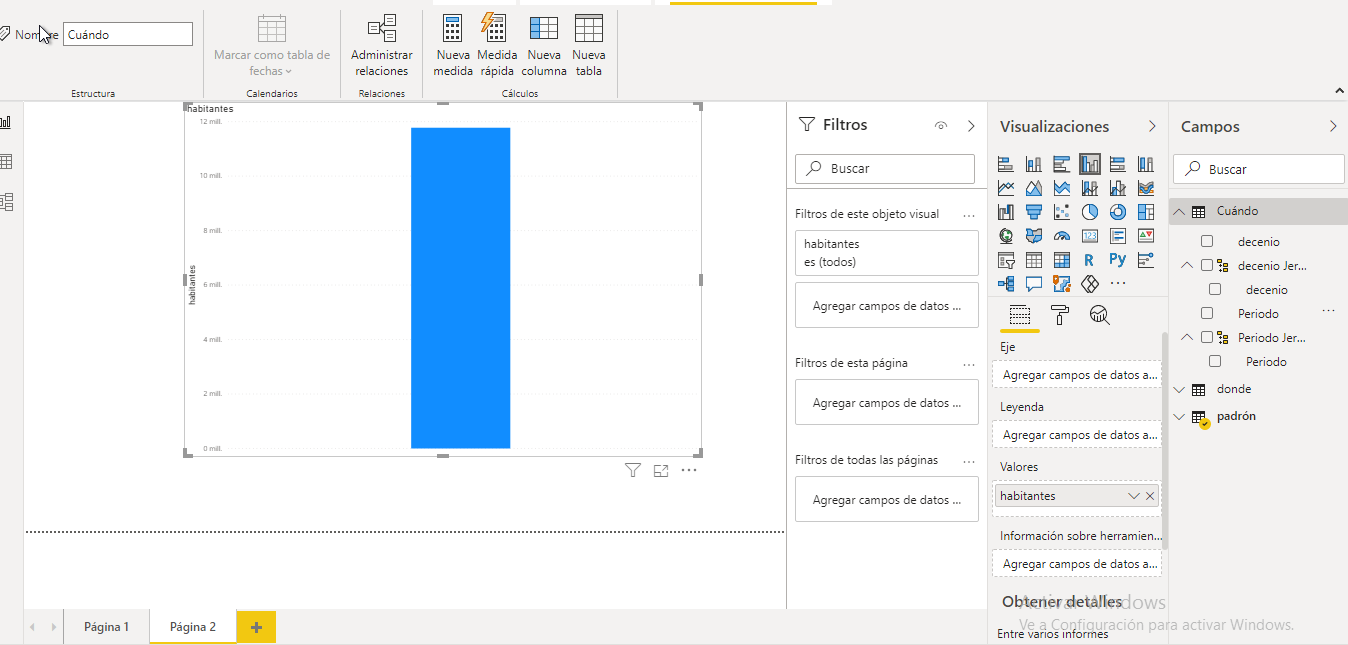
Añadimos a esta nueva jerarquía los campos “comarca” y “provincia” pinchando sobre los 3 puntos de la derecha del campo y “añadir a jerarquía”.



### Ejercicio 4.2

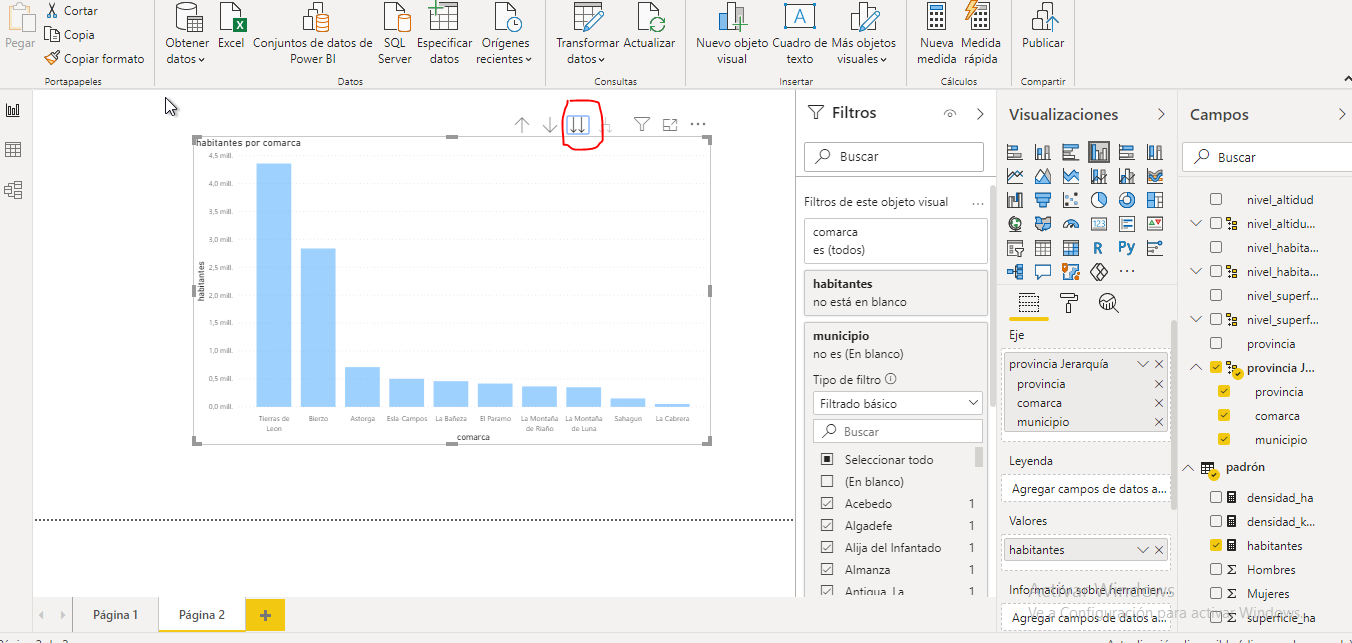
Define las jerarquías asociadas a los campos provincia, comarca, municipio, nivel\_altitud, nivel\_habitantes y nivel\_superficie según las explicaciones previas. Define también la jerarquía asociada a los campos decenio, y periodo (captura una pantalla donde se muestren todas las jerarquías).



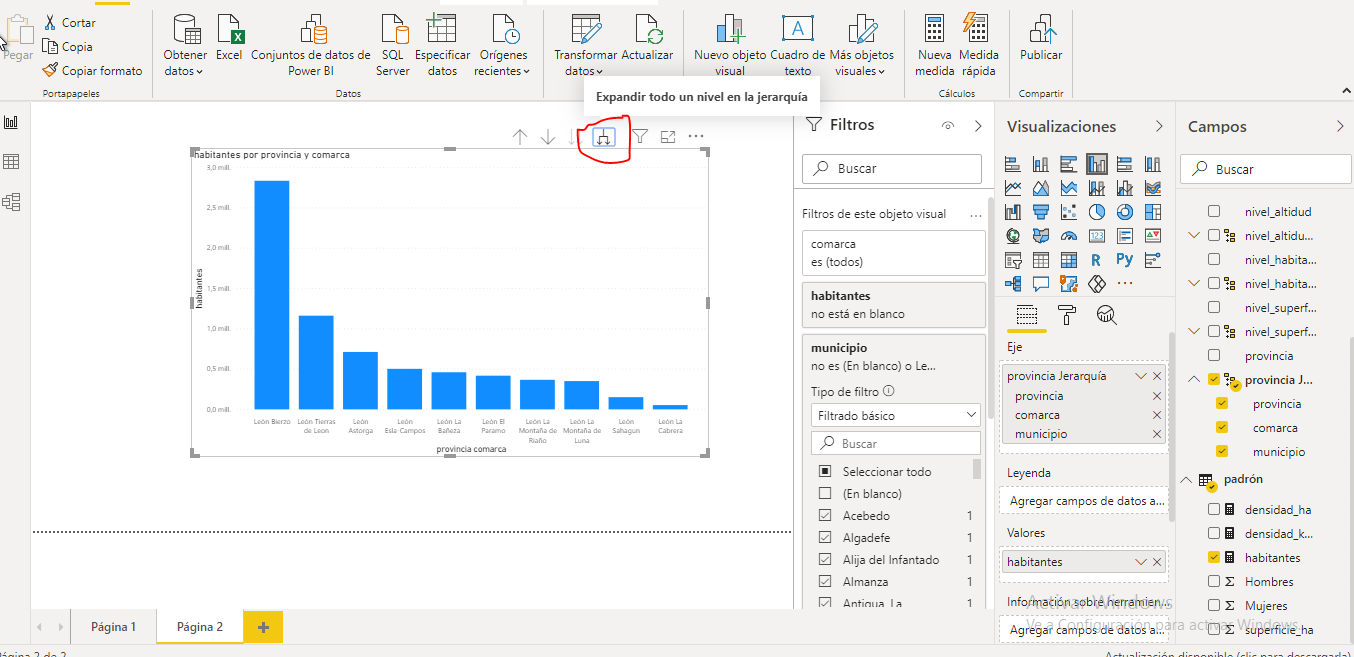


## Consultas basadas en jerarquías

Al incorporar una jerarquía al informe, se añaden todos los campos a los filtros. Pinchando encima del gráfico en las dos flechitas hacia abajo, vamos al siguiente nivel de la jerarquía.



Para tener un mayor detalle, pincharemos en la flecha hacia abajo y podemos expandir toda la jerarquía en el nivel más alto pinchando en la flecha que se desglosa.



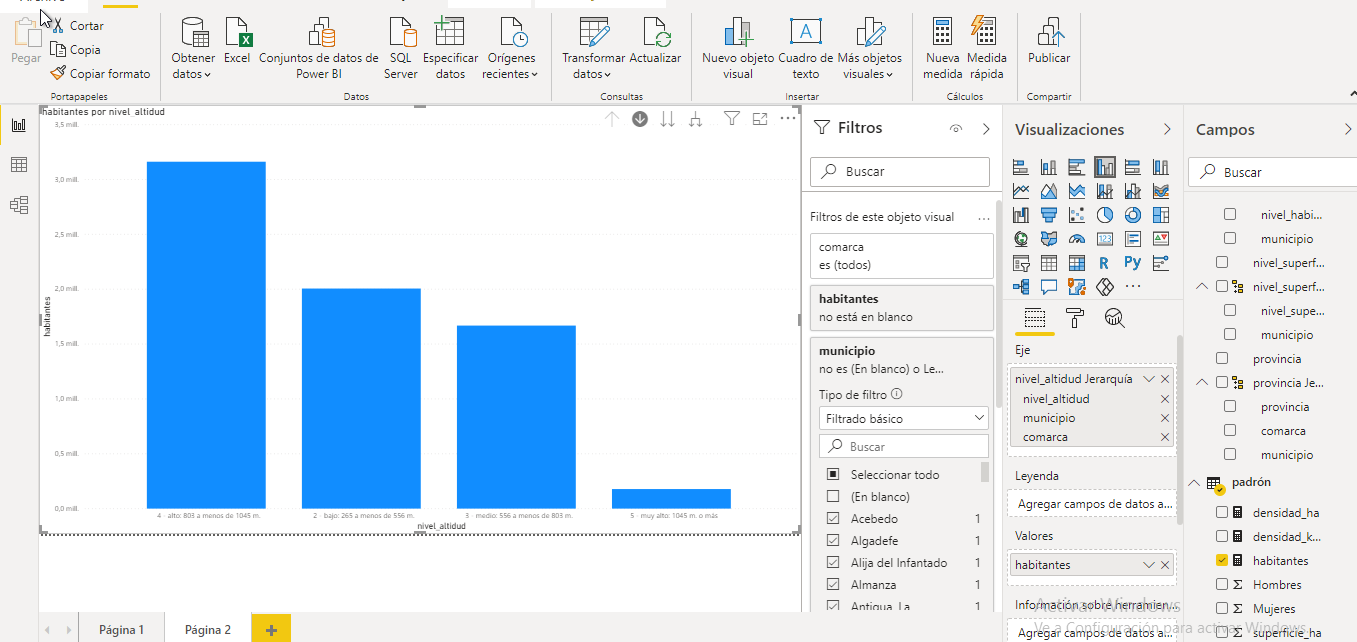
### Ejercicio 4.3

Partiendo de un informe inicial cualquiera (se puede obtener aplicando las operaciones que se consideren necesarias sin necesidad de explicarlas), genera nuevos informes mediante las operaciones multidimensionales. Una vez se tenga el informe inicial, el resto de informes han de generarse **de forma consecutiva** (sin saltarse informes intermedios).

* Para las operaciones Roll-Up y Drill-Down usa jerarquías utilizando al menos una vez cada una de las operaciones específicas sobre jerarquías.
* Para cada informe obtenido, indica la **operación multidimensional aplicada**y el **nivel de detalle** de los datos que representa (captura una pantalla para cada informe).

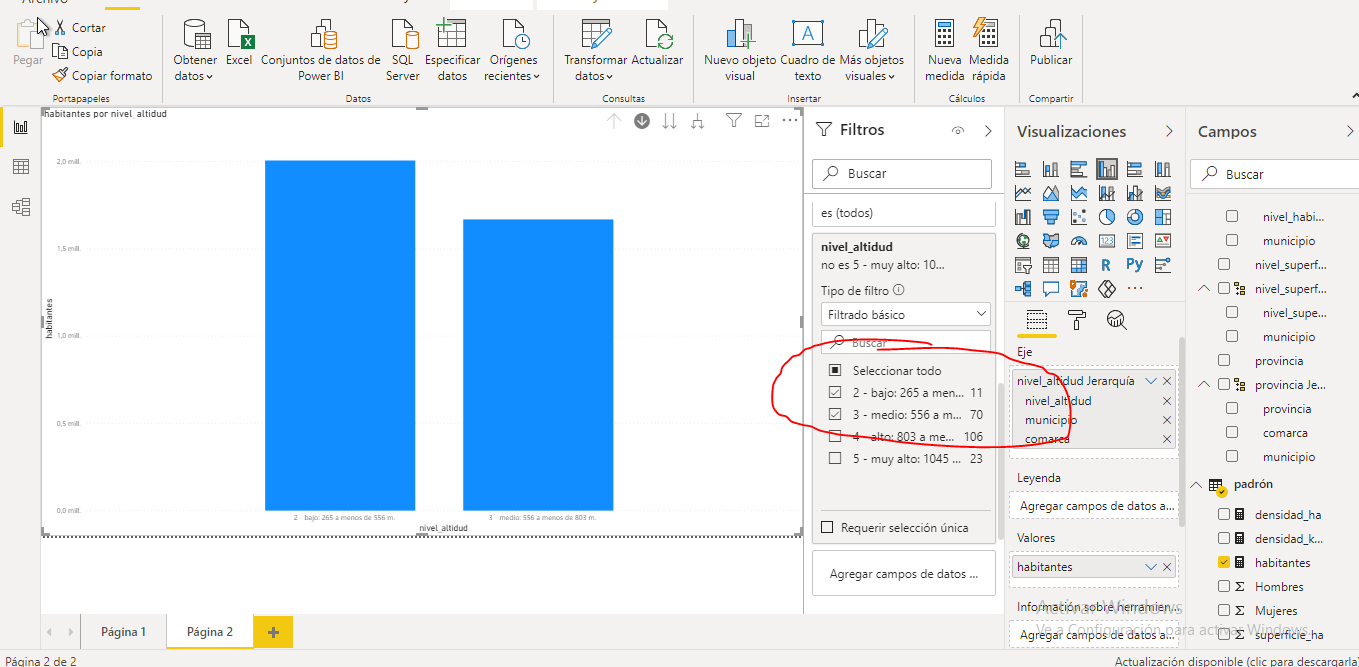
Parto de una nueva jerarquía(añadida a nivel\_altitud el campo comarca).

En este caso, no hay ninguna jerarquía “válida” para este ejercicio. Solamente podría utilizarse para ver los niveles de jerarquía la jerarquía utilizada anteriormente “provincia Jerarquia”, ya que podemos observar el número de habitantes a un nivel de provincia inicialmente, luego a un nivel de municipio y finalmente, a un nivel más detallado, a nivel de comarca. De menor a mayor detalle.



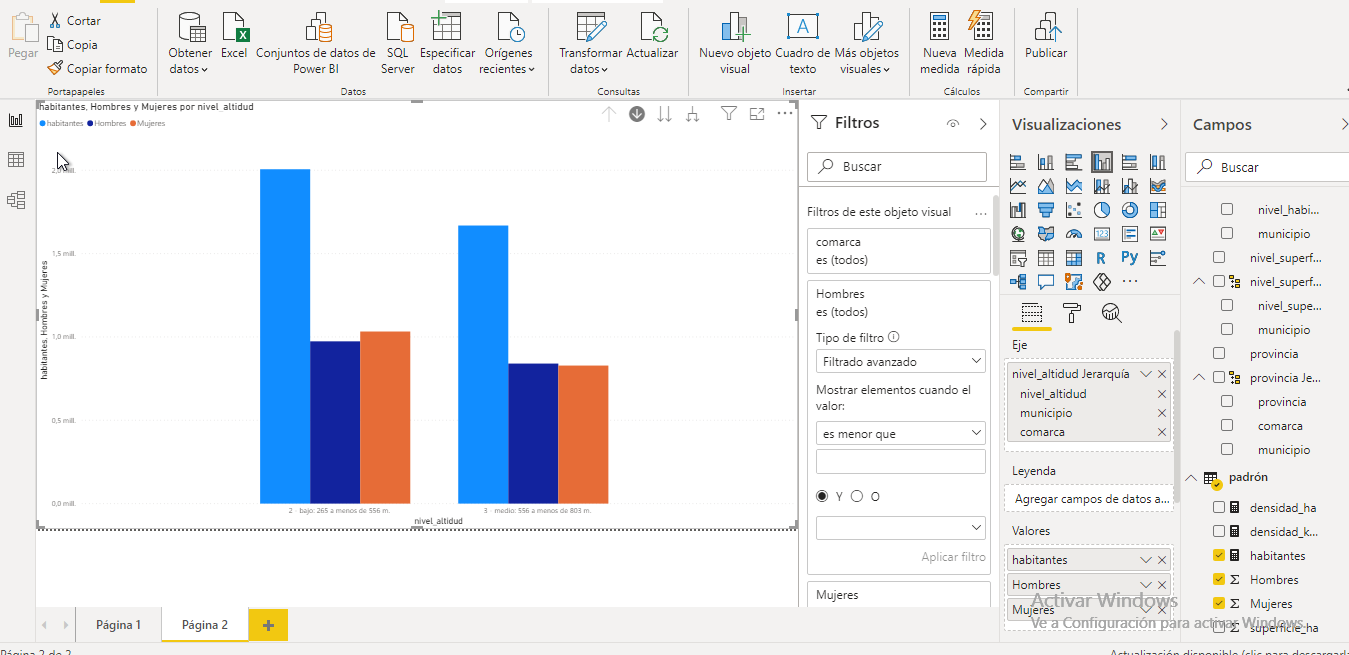
Slice&Dice

Partiendo del informe inicial, seleccionamos la cantidad de habitantes que tienen sus residencias a menos de 803 metros(nivel 4)



Drill-Down

Añadimos hombres y mujeres, para un informe más fino. Podemos observar la cantidad de hombres y mujeres y la sumatoria de ambos por nivel de altitud, menor de 803 metros.



Roll-Up

Eliminaremos la columna de sumatorio de habitantes, ya que no nos hace falta realmente.

