



**Universitat Oberta
de Catalunya**

Máster universitario de Ciencia de Datos

Práctica 3

**Diseño y uso de bases de datos analíticas – Explotación
de datos.**

Autor:

Mario Ubierna San Mamés

Índice de Contenido

Índice de Contenido	3
Índice de tablas	4
Índice de ilustraciones	5
1. Introducción	6
1.1. Presentación	6
1.2. Descripción	6
2. Creación del modelo OLAP	8
2.1. Creación de la estructura física del modelo	8
2.2. Creación del proyecto.....	8

Índice de tablas

No se encuentran elementos de tabla de ilustraciones.

Índice de ilustraciones

No se encuentran elementos de tabla de ilustraciones.

1.Introducción

1.1. Presentación

A partir de la solución oficial de la segunda práctica (PRA2) el estudiante debe implementar los cubos multidimensionales necesarios para la explotación de la información y el posterior análisis de datos. De este modo se facilitará la toma de decisiones de los usuarios potenciales.

Así pues, esta actividad tiene el objetivo de implementar un modelo multidimensional online analytical processing (OLAP) para el análisis multidimensional de la información con el fin de responder a las preguntas definidas en el análisis de requerimientos.

Adicionalmente, se facilitará junto a este enunciado el fichero «*export_DW_COVID.sql*», que contiene los scripts de generación y carga de todas las tablas planteadas en la solución, para que el estudiantado pueda partir de la misma base.

1.2. Descripción

Más concretamente esta tercera parte del caso práctico consiste en diseñar un modelo OLAP para el análisis multidimensional de la información disponible en el almacén de datos que permita dar respuesta a las siguientes cuestiones:

- Análisis de las provincias con mayor porcentaje de movilidad según datos móviles.
- Análisis del porcentaje de la población que evitaba las aglomeraciones según la comunidad autónoma.
- Análisis del promedio de sanciones por habitante.
- Evolución de las llamadas de urgencia al 112 en Cataluña por tipología de llamada.

- Análisis de las llamadas de urgencia frente al porcentaje de la población que evitaba las aglomeraciones entre los meses comprendidos entre marzo y junio de 2020 en Cataluña, desglosado por provincia.
- Determinación del día de la semana con menor número de denuncias.
- Análisis de las diez fechas (*top ten*) con mayor número de llamadas de urgencia al 112 con tipología de tránsito registrada, tanto en época de *COVID* como antes.

2. Creación del modelo OLAP

En este apartado vamos a crear tanto el proyecto en *Visual Studio*, como observar el origen de datos, la creación de los cubos y la decisión sobre jerarquías/dimensiones/atributos relacionados.

2.1. Creación de la estructura física del modelo

Lo primero que debemos de hacer es ejecutar el *script* que se nos ha proporcionado “*export_DW_COVID.sql*”, ya que haciendo esto obtenemos el diseño tanto de los hechos como de las dimensiones junto a sus datos.

Cabe destacar que para ejecutar el *script* lo hemos tenido que modificar añadiendo las dos primeras líneas siguientes, para que así haga uso de la base de datos correcta:

```
USE [DB_marioum]  
GO
```

Ilustración 1 - Base de datos DB_marioum.

2.2. Creación del proyecto

Una vez que tenemos todos los datos que vamos a hacer uso de ellos para explotarlos, podemos crear el proyecto en *Visual Studio*.

Para ello abrimos el programa y creamos un nuevo proyecto de tipo “*Analysis Services Multidimensional*”:

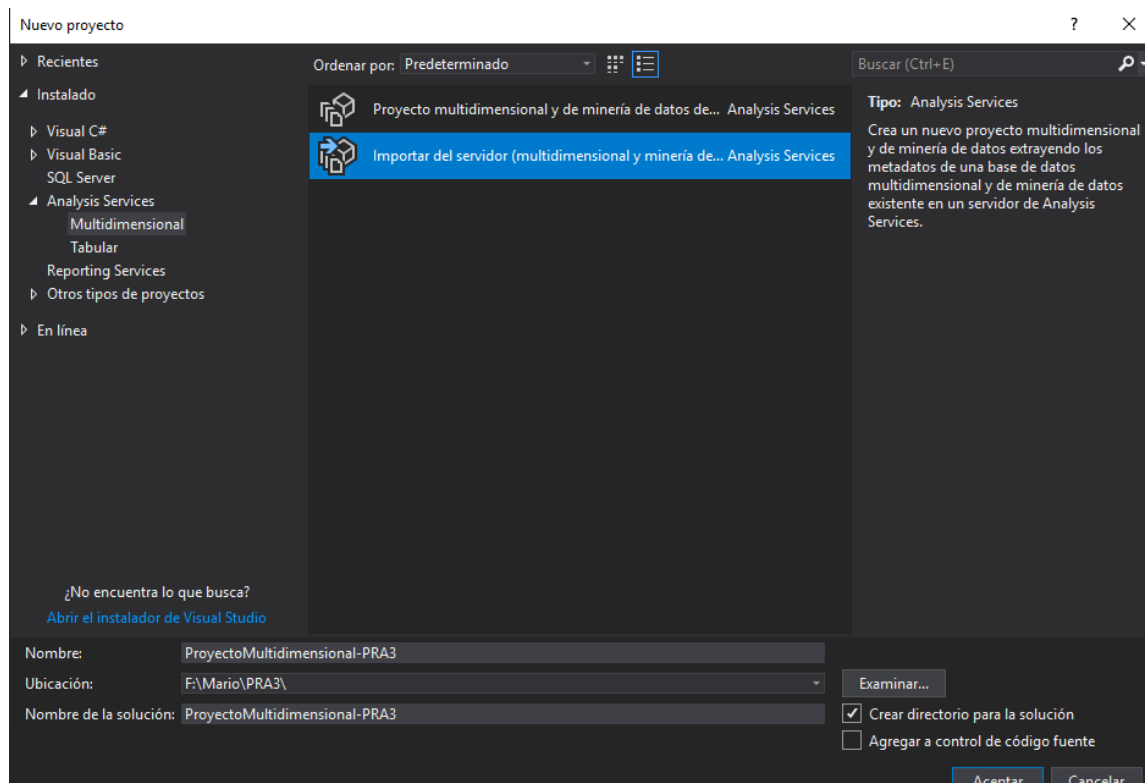


Ilustración 2 - Creación del proyecto multidimensional.

Una vez que hemos seleccionado la opción “Importar del servidor” y establecemos el nombre del proyecto junto con su ubicación, nos aparecerá un asistente para configurar la base de datos a usar.

Para ello indicamos el servidor y en el menú desplegable de la base de datos, seleccionamos la nuestra, es decir, “DB_mariouism”:

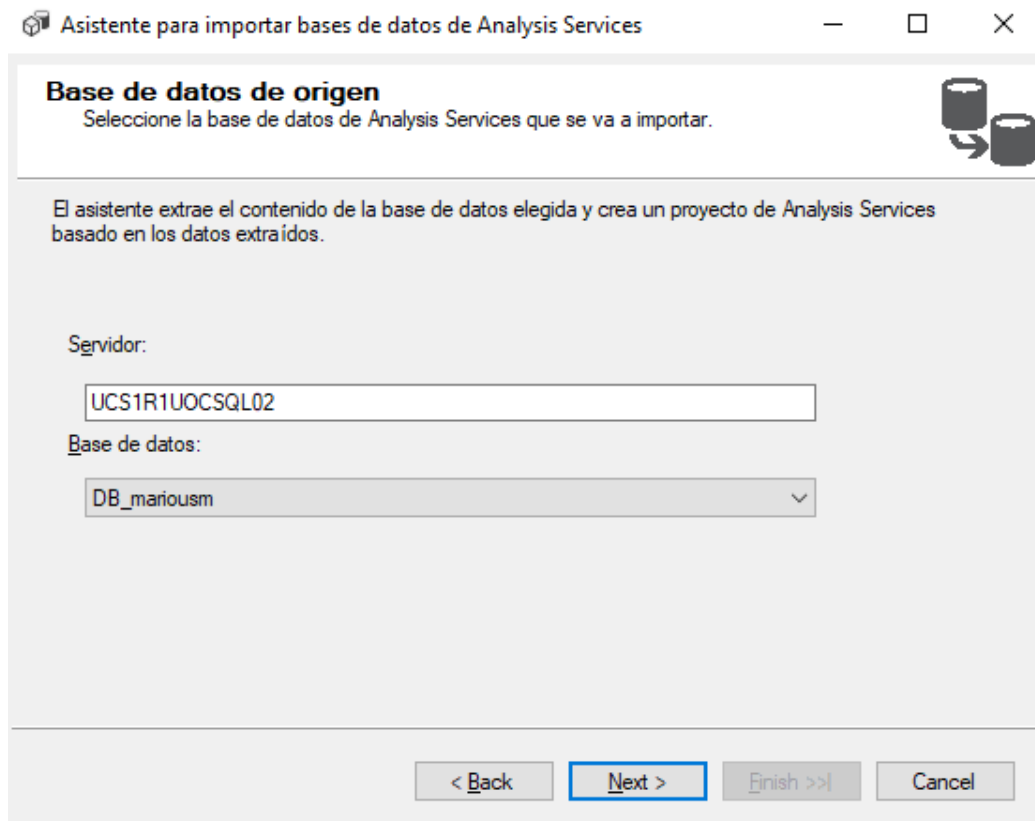


Ilustración 3 - Configuración origen de los datos.

Una vez importada la estructura de la base de datos a nuestro proyecto, hay que configurar el destino de los datos. Para ello, nos dirigimos a las propiedades del proyecto:

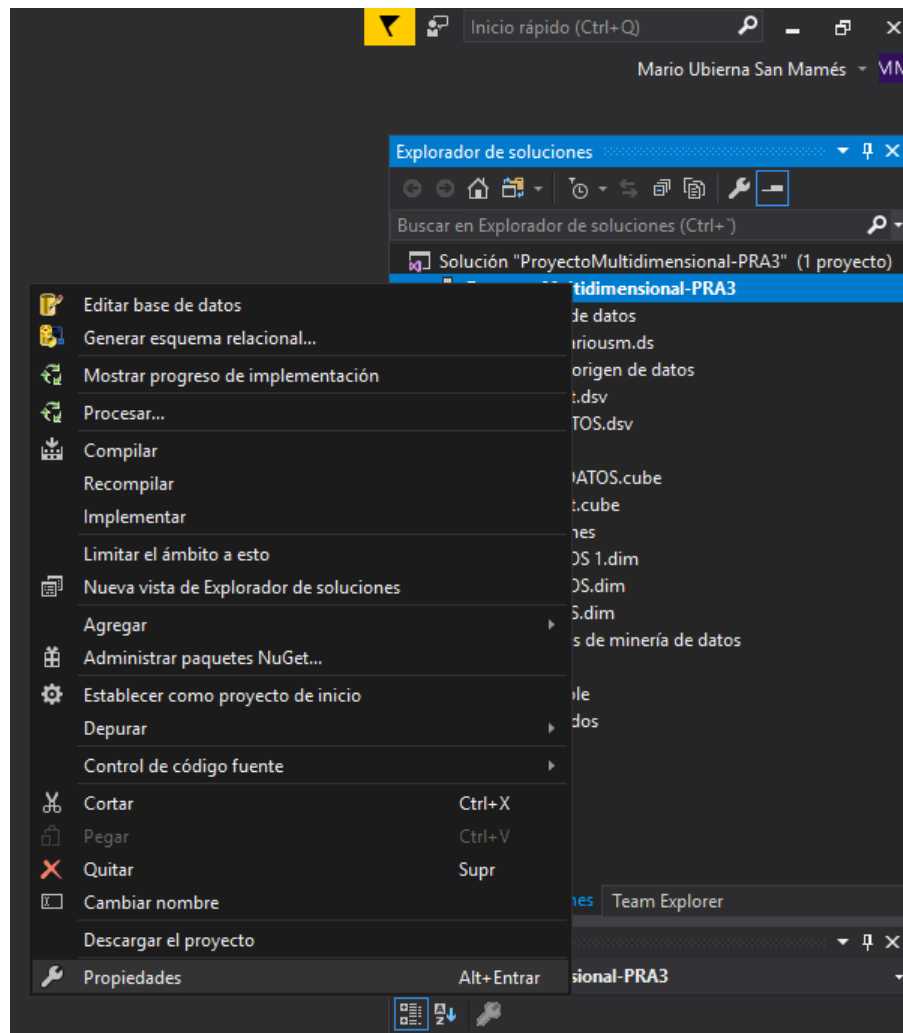


Ilustración 4 - Propiedades del proyecto Visual Studio.

En las propiedades nos vamos al apartado implementación, y en él cambiamos el nombre del servidor tal y como se muestra en la siguiente captura:

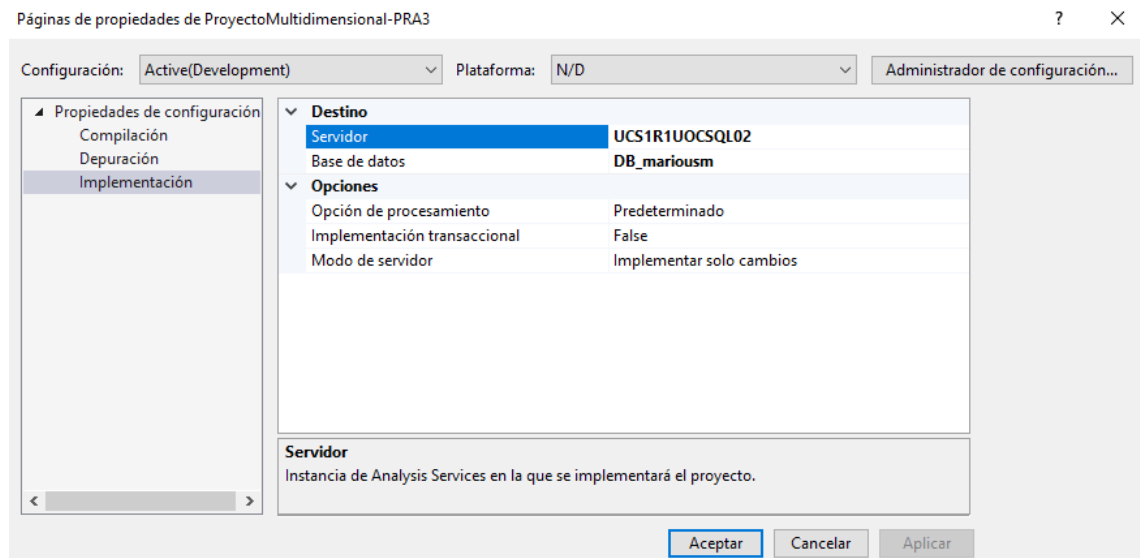


Ilustración 5 - Configuración destino de datos en Visual Studio.

Finalmente hay que configurar el origen de los datos, para ello sobre el nombre de la base de datos en el explorador de soluciones, hacemos doble click y nos mostrará la siguiente ventana:

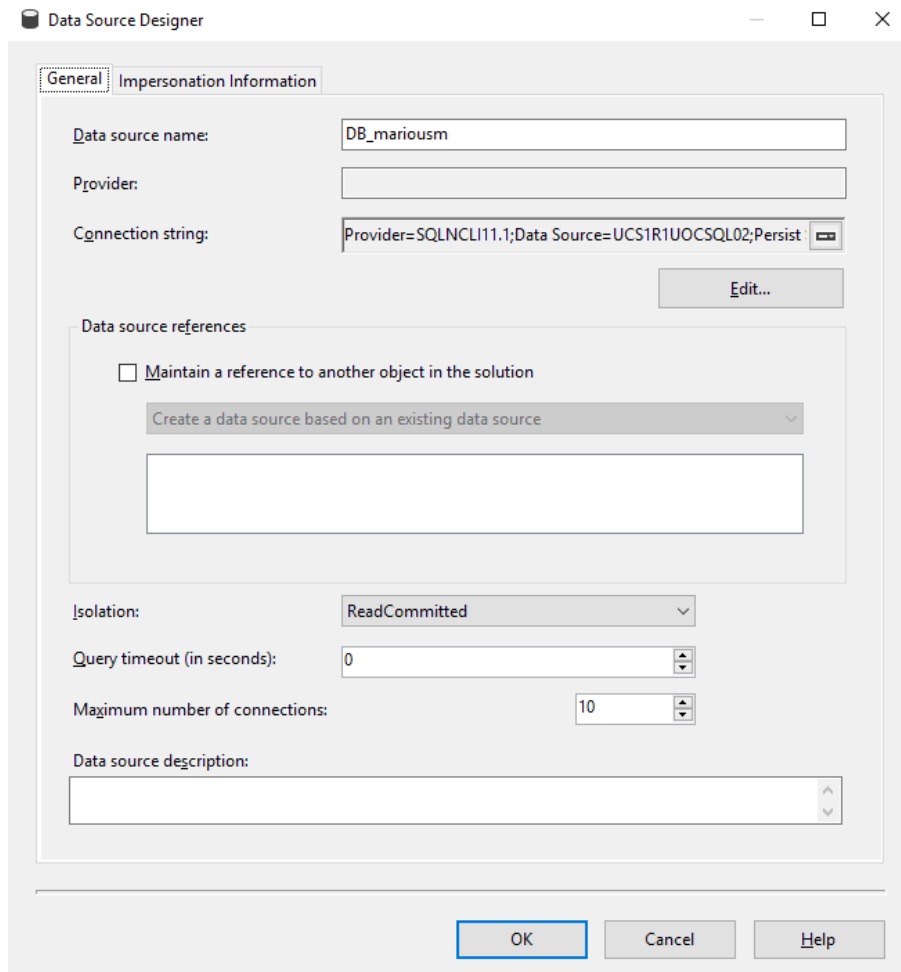


Ilustración 6 - Ventana origen de datos.

Pulsamos sobre el botón “*Edit*” para cambiar la conexión a la base de datos, una vez hemos pulsado veremos la siguiente pantalla:

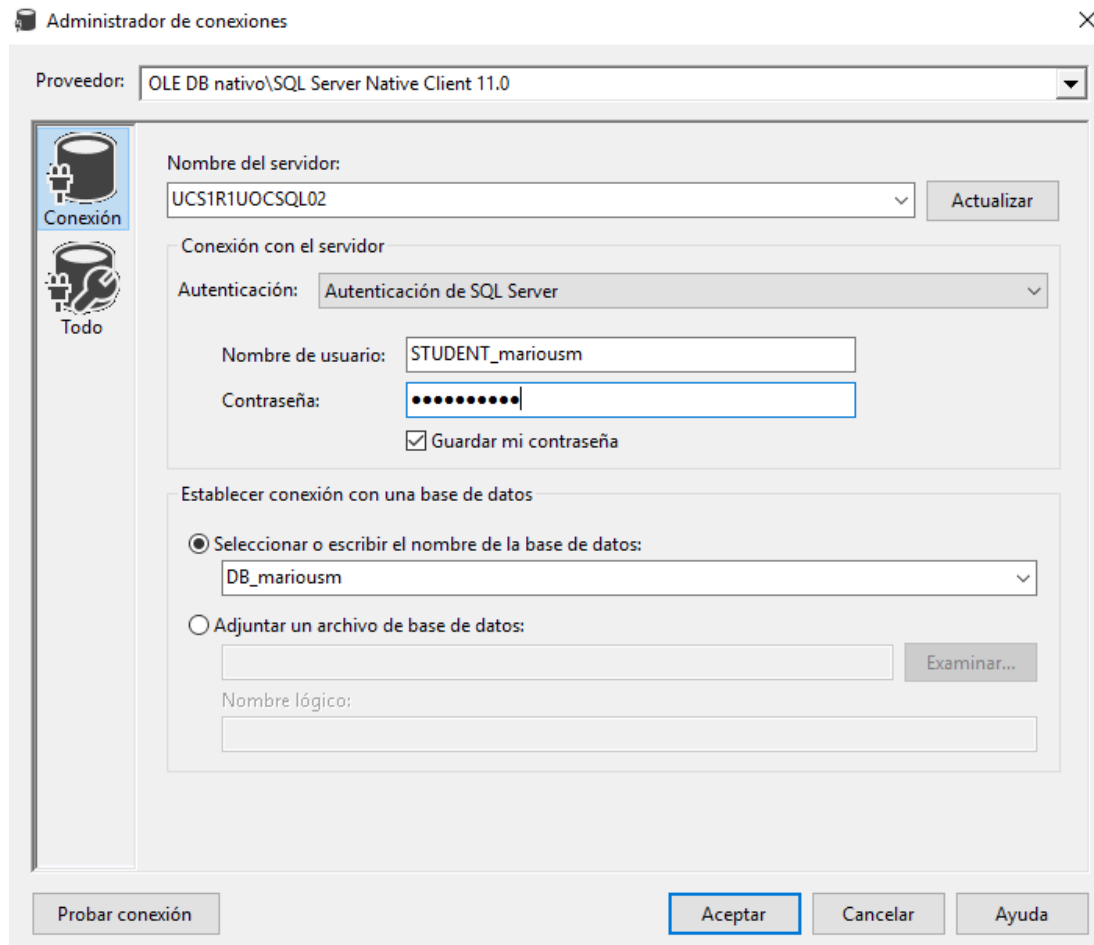


Ilustración 7 - Modificación conexión de datos.

Introducimos el servidor que se nos ha proporcionado, el nombre del usuario y la contraseña que usamos para acceder a *Microsoft SQL Server Management Studio 2017*.

Una vez realizados todos estos pasos podemos dar por concluido este apartado.

2.3. Vistas del origen de datos

Una vez que ya tenemos configurado nuestro proyecto, podemos comenzar a definir las diferentes vistas respecto al origen de los datos.

En nuestro caso vamos a tener dos:

- Vista de Llamadas112: nos va a permitir realizar el acceso a la tabla correspondiente de las llamadas y así hacer un análisis de las mismas.
- Vista de Mediciones: en ella vamos a poder realizar el acceso a la tabla de métricas y hacer un profundo análisis de las mismas.

Para crear las dos vistas debemos acceder al explorador de soluciones del proyecto y hacer click derecho sobre la carpeta “Vistas del origen de datos”, posteriormente pulsaremos sobre la opción “Nueva vista del origen de datos”:

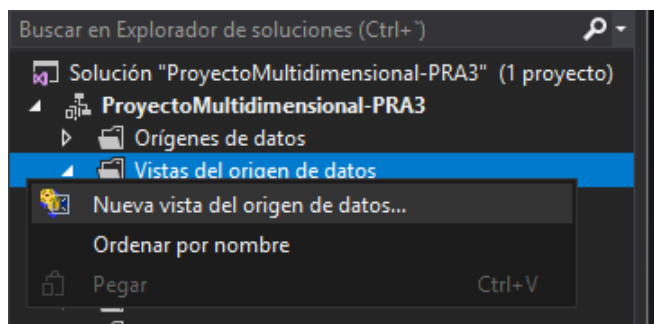


Ilustración 8 - Creación de nueva vista del origen de datos.

Una vez que hemos accedido al asistente para la creación de la vista, nos aparecerá la siguiente ventana en la que tenemos que seleccionar el origen de datos:

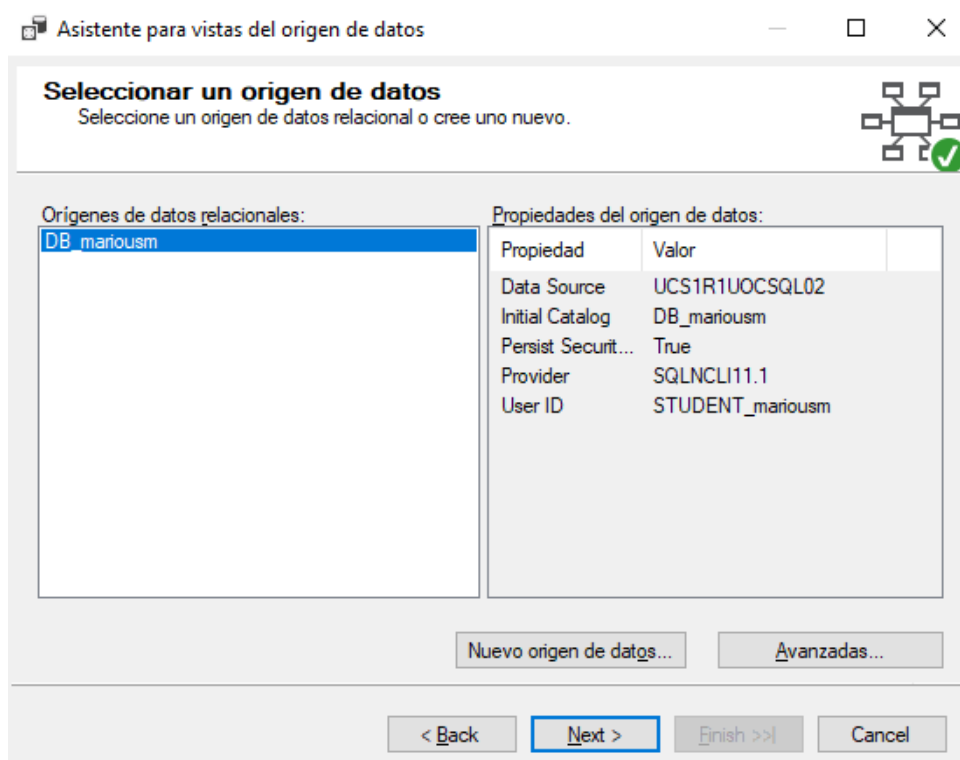


Ilustración 9 - Selección origen de datos vista.

2.3.1. Vista Llamadas112

Posteriormente seleccionamos tanto la tabla del hecho como de las dimensiones correspondientes, en nuestro caso el hecho es “FACT_Llamadas112” y las dimensiones: “DIM_Ambito_Geografico”, “DIM_Fecha”, “DIM_Tipologia”:

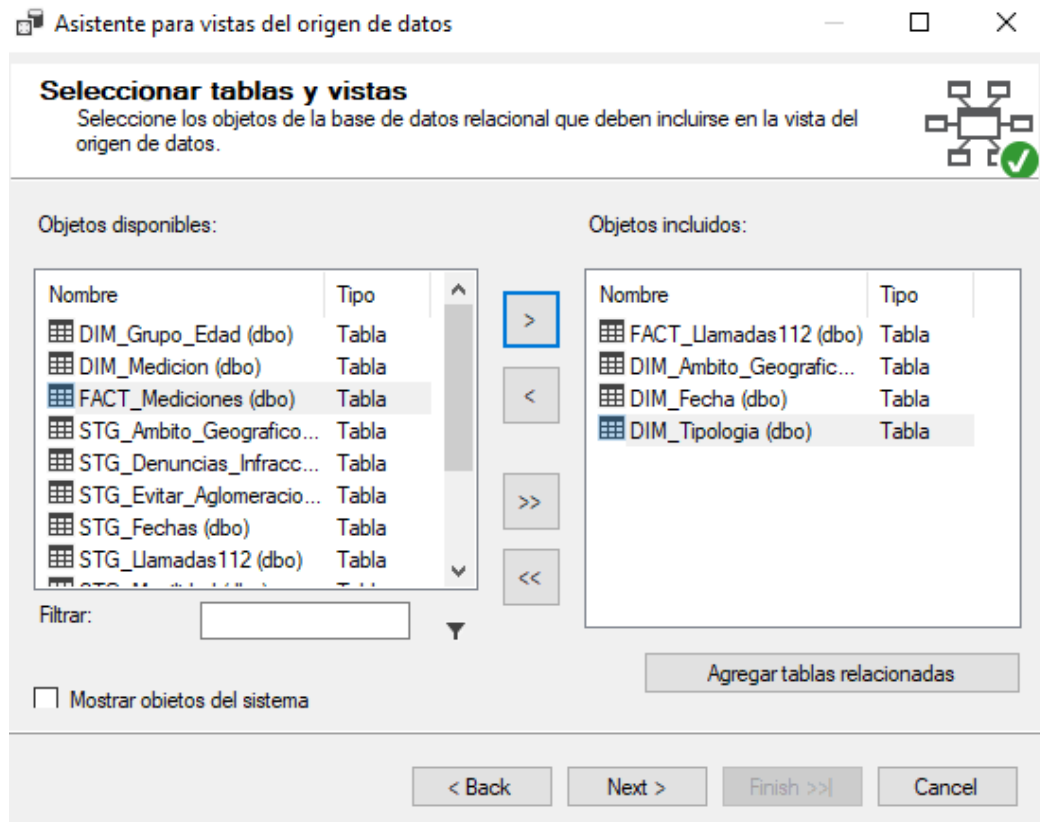


Ilustración 10 - Selección tablas para la vista Llamadas112.

Pulsamos el botón “Next”, indicamos el nombre de la vista “vLlamadas112” y finalizamos el asistente.

Al hacer doble click sobre la vista nos muestra el diseño de la misma:

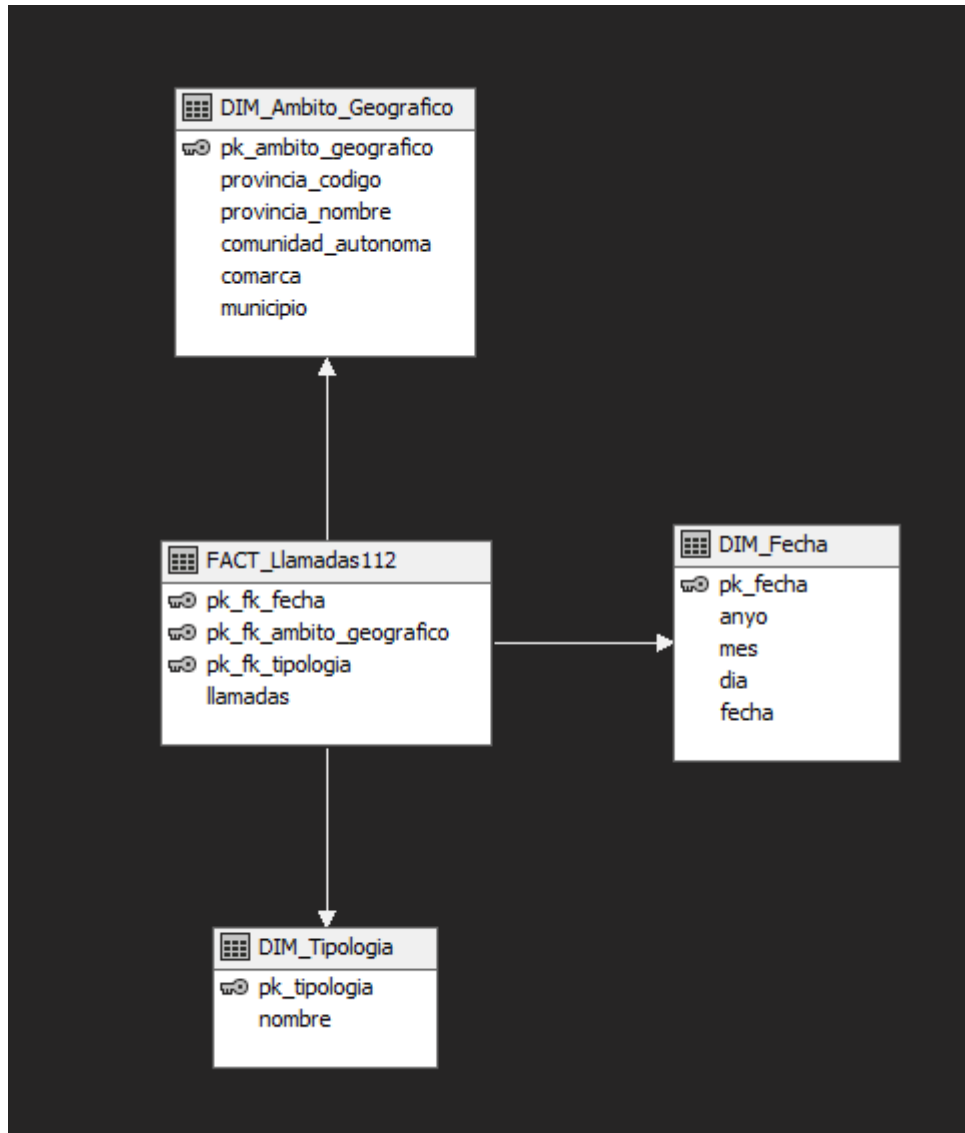


Ilustración 11 - Diseño vista vLlamadas112.

2.3.2. Vista Métricas

A continuación, vamos a realizar el mismo proceso que el detallado en la creación de la vista anterior, pero en este caso vamos a crear una vista con el fin de explotar la información sobre las denuncias, el porcentaje de población que evitaba aglomeraciones...

Esta vista tiene en cuenta el hecho “FACT_Mediciones” y las dimensiones: “DIM_Ambito_Geografico”, “DIM_Fecha”, “DIM_Grupo_Edad” y “DIM_Medicion”:

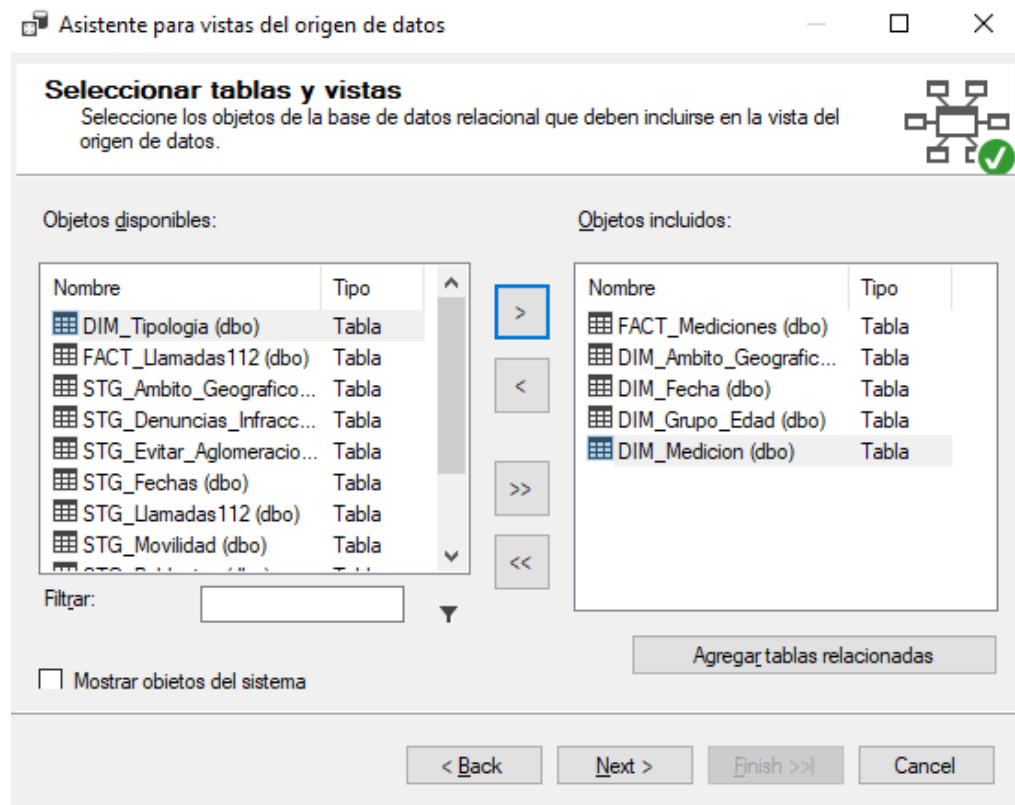


Ilustración 12 - Selección de tablas para la vista Mediciones.

Pulsamos el botón “Next”, indicamos el nombre de la vista “vMediciones” y finalizamos el asistente:

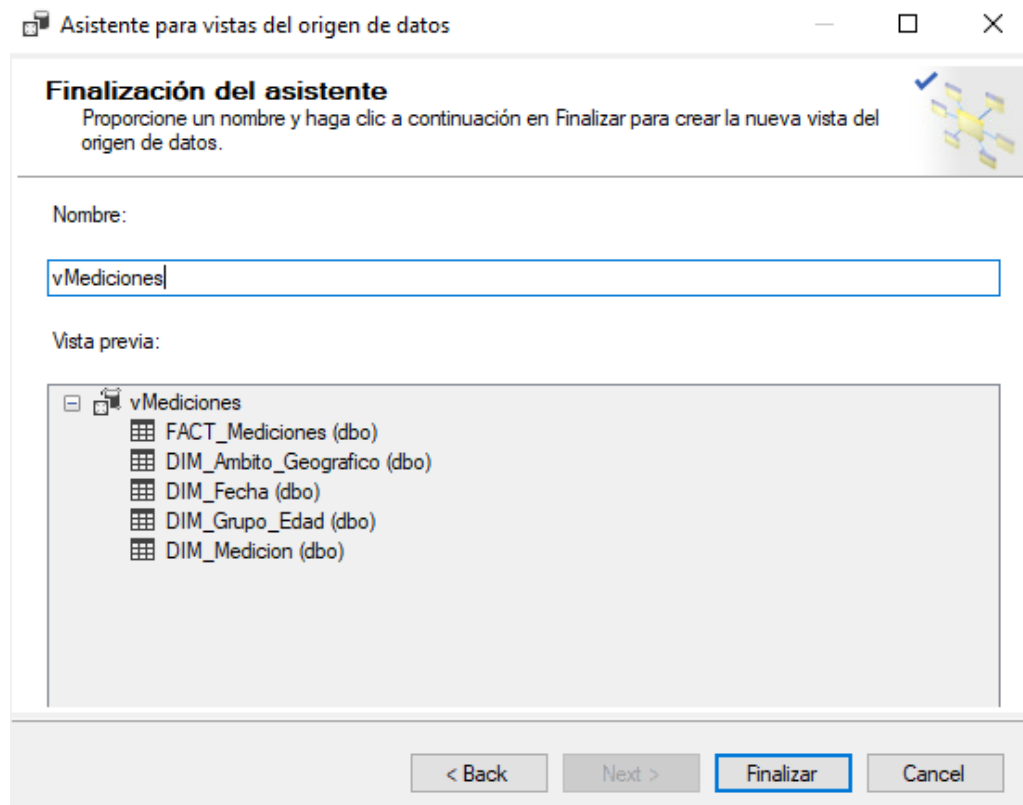


Ilustración 13 - Nombre de la vista vMediciones.

Al hacer doble click sobre la vista nos muestra el diseño de la misma, comprobamos que todo está de forma correcta:

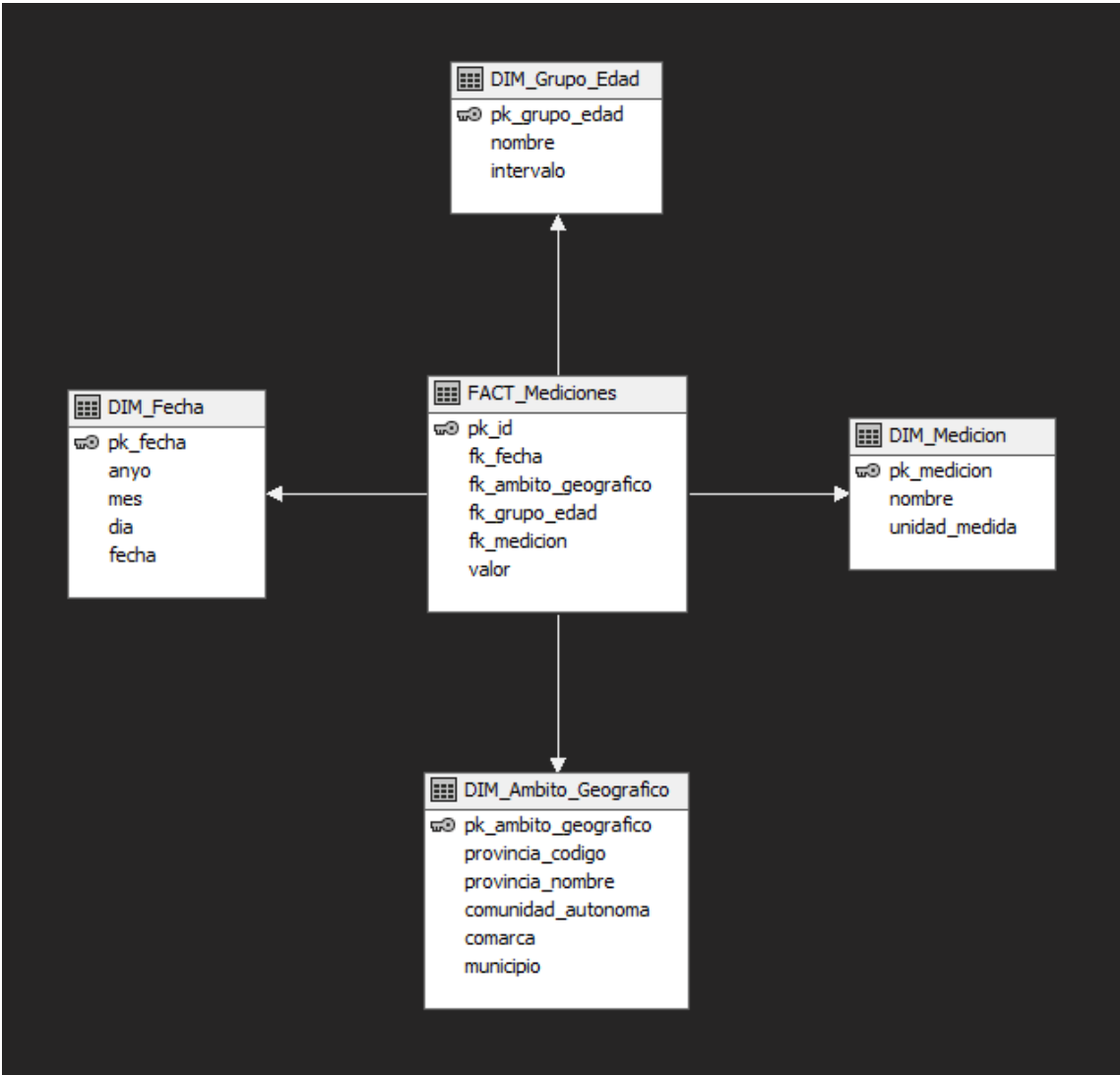


Ilustración 14 - Diseño vista vMediciones.

2.4. Creación de cubos