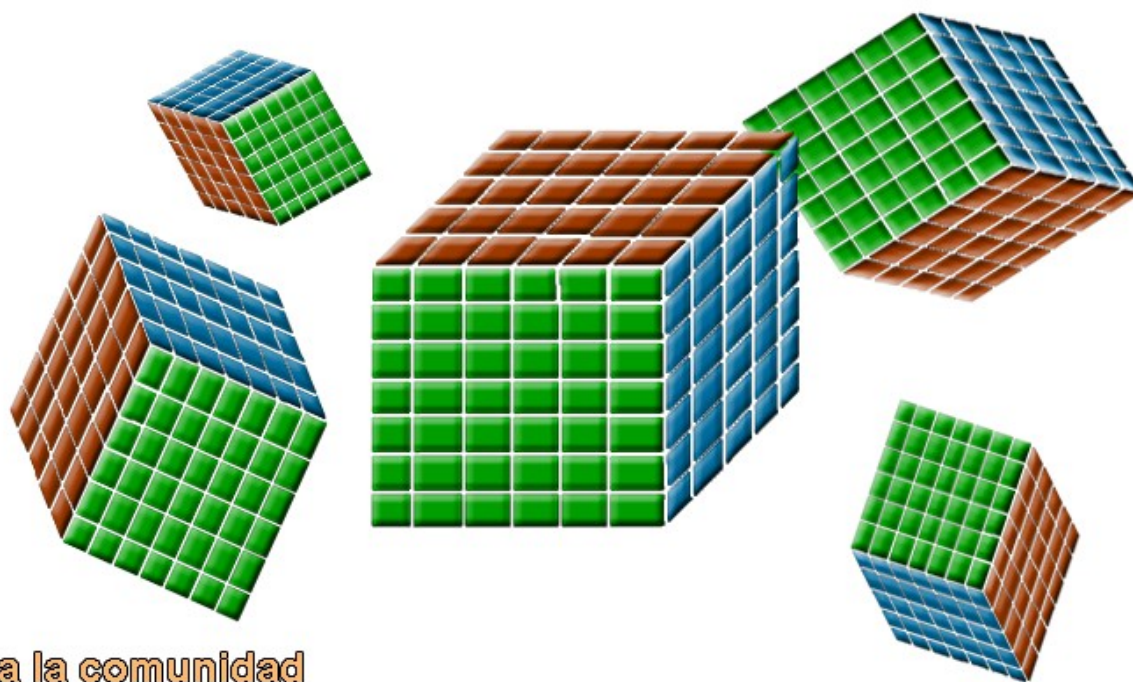


“ Pasos para crear Cubos con Mondrian Schema Workbench”

Por : Ing. Dennis Alba Infante



Para la comunidad

Open Business Intelligence



La red del Business Intelligence

Copyright 2009 Ing. Dennis Alba Infante. Se otorga permiso para copiar, distribuir y/o modificar este documento bajo los términos de la Licencia de Documentación Libre de GNU, Versión 1.3 o cualquier otra versión posterior publicada por la Free Software Foundation. Una copia de la licencia se puede encontrar en el siguiente enlace: <http://www.gnu.org/licenses/fdl-1.3.html>

Fecha	Versión	Autor	Detalle del cambio
03/11/2009	0.1	Ing. Dennis Alba Infante	Versión Inicial.
06/11/2009	0.2	Ing. Bernabeu Ricardo Dario	Revisión.

Para ver la traducción no oficial de la Licencia de Documentación Libre de GNU, seguir el siguiente enlace: http://stuff.danexnow.org/gfdl_es.html

Pasos para crear Cubos con Mondrian Schema Workbench

- [1\) Instalación](#)
- [2\) Configuración](#)
- [3\) Acceso](#)
- [4\) Conexión](#)
- [5\) Nuevo Schema](#)
- [6\) Nuevo Cubo](#)
- [7\) Definir Tabla de Hechos](#)
- [8\) Añadir Dimensión “Tipos”](#)
- [9\) Añadir Dimensión “Fechas”](#)
- [9\) Añadir Dimensión “Fechas”](#)
- [11\) Guardar el Cubo](#)
- [12\) Publicar el Cubo](#)
- [13\) Ejemplos de Análisis Interactivos](#)

1) Instalación:

Lo primero que debemos hacer es descargar Mondrian Schema Workbench (a.k.a Pentaho Schema Workbench -PSW-) de la siguiente dirección:

- <http://sourceforge.net/projects/mondrian/files/>

La última versión al día de hoy es la 3.1.1-estable:

- En Window\$: seleccionar el archivo psw-ce-3.1.1.12687.zip
- En GNU/Linux: seleccionar el archivo psw-ce-3.1.1.12687.tar.gz

Una vez que se descargó el archivo, procedemos a descomprimirlo en algún lugar apropiado.

2) Configuración:

Una de las primeras cosas que se debe configurar al utilizar PSW, es la ubicación de los .jar que se emplearán en las conexiones con las distintas Bases de Datos.

Para el caso de ORACLE es "ojdbc14.jar", recuerden que lo deben ubicar en la carpeta "drivers" que se encuentra en la carpeta raíz en donde se instaló PSW (la carpeta por defecto se llama "schema-workbench").

3) Acceso:

Luego se debe iniciar PSW:

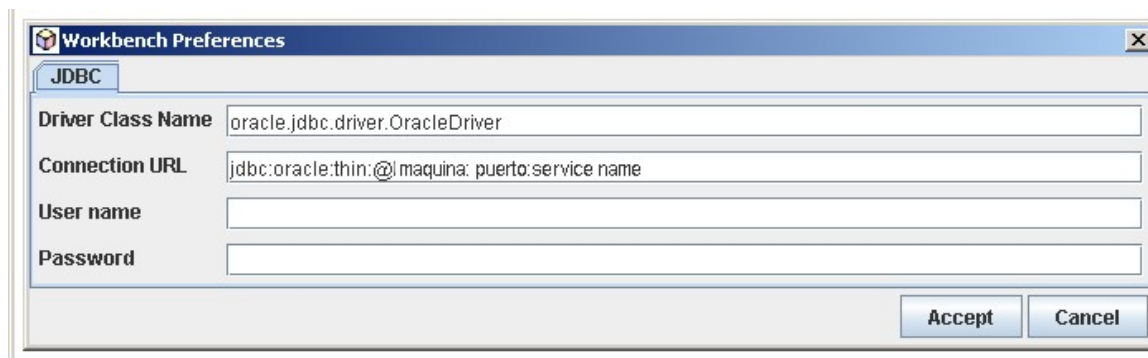
- En Window\$: ir al directorio raíz y ejecutar "workbench.bat"
- En GNU/Linux: abrir una consola, ir al directorio raíz y escribir:
 - `chmod +x *.sh`
 - `sh workbench.sh`

4) Conexión:

Una vez que ingresamos a PSW, se podrá establecer la conexión.

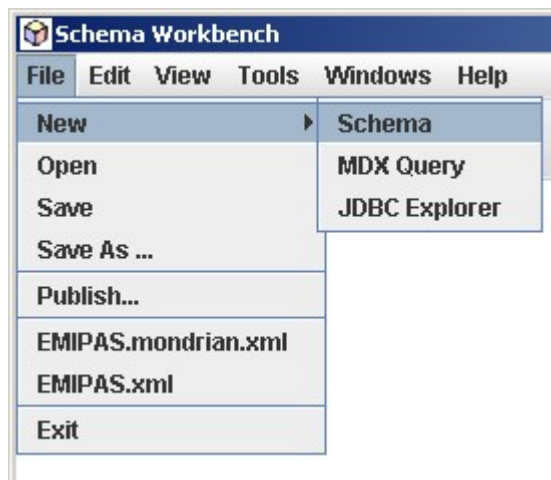


Llenan los campos:

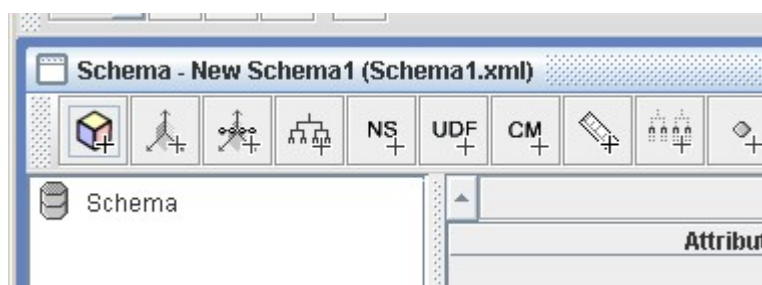


5) Nuevo Schema:

Creamos el esquema nuevo: seleccionar en el menú “File → New → Schema” (esperen que se toma su tiempo)



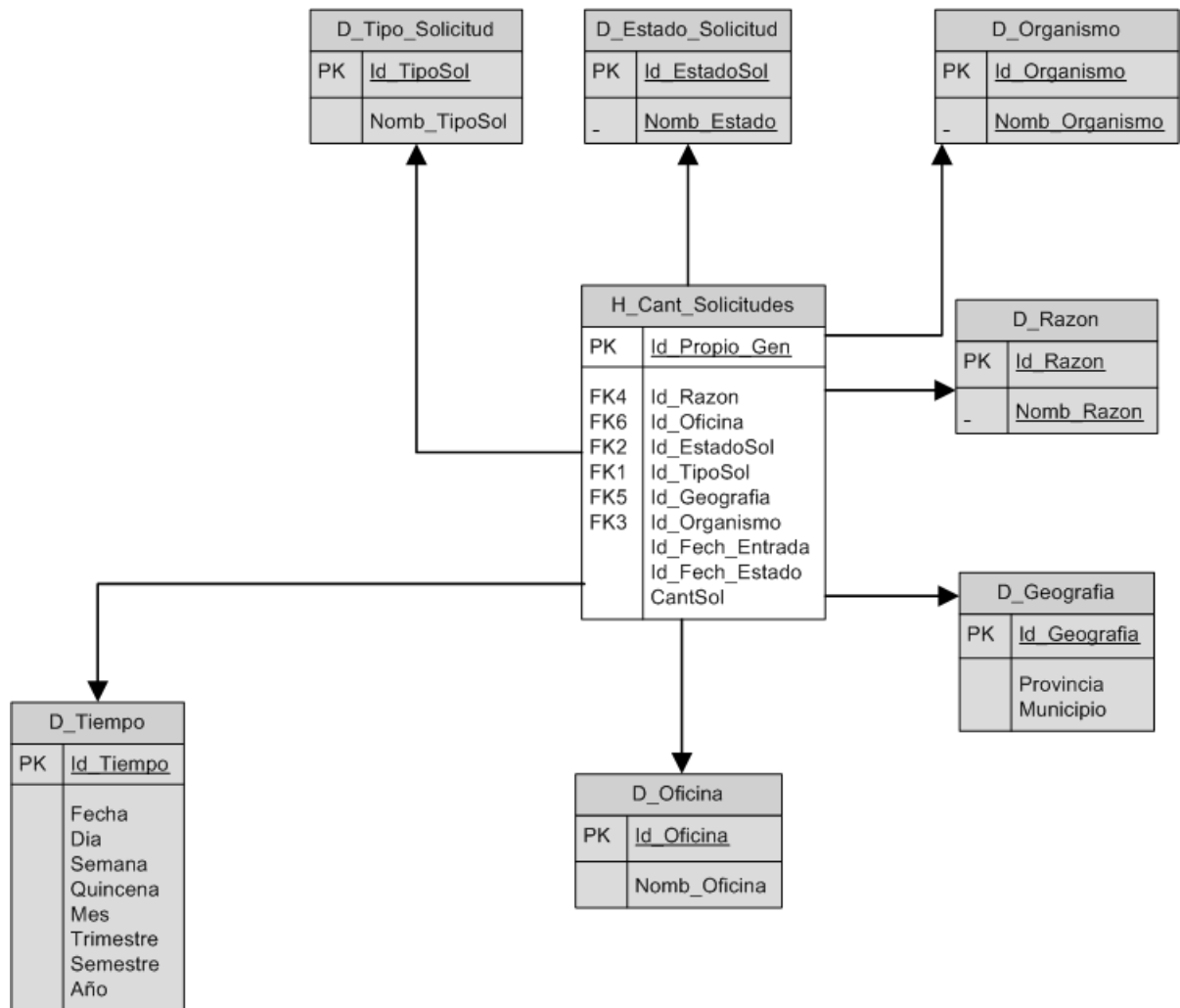
Cuando hayamos conectado con la base de datos veremos que se queda en blanco, como si nada hubiera ocurrido:



Pero si ocurrió, ya está todo listo para crear nuestro cubo.

Partamos de un esquema en estrella que tengamos en nuestro Data Warehouse, de dicha estrella es que modelaremos el cubo.

En el sistema de ejemplo se recogen solicitudes de diversas empresas para que se les repare o dé mantenimientos a sus respectivas flotas de autos y existen varios tipos de solicitudes en dependencia de la empresa que solicita el servicio.

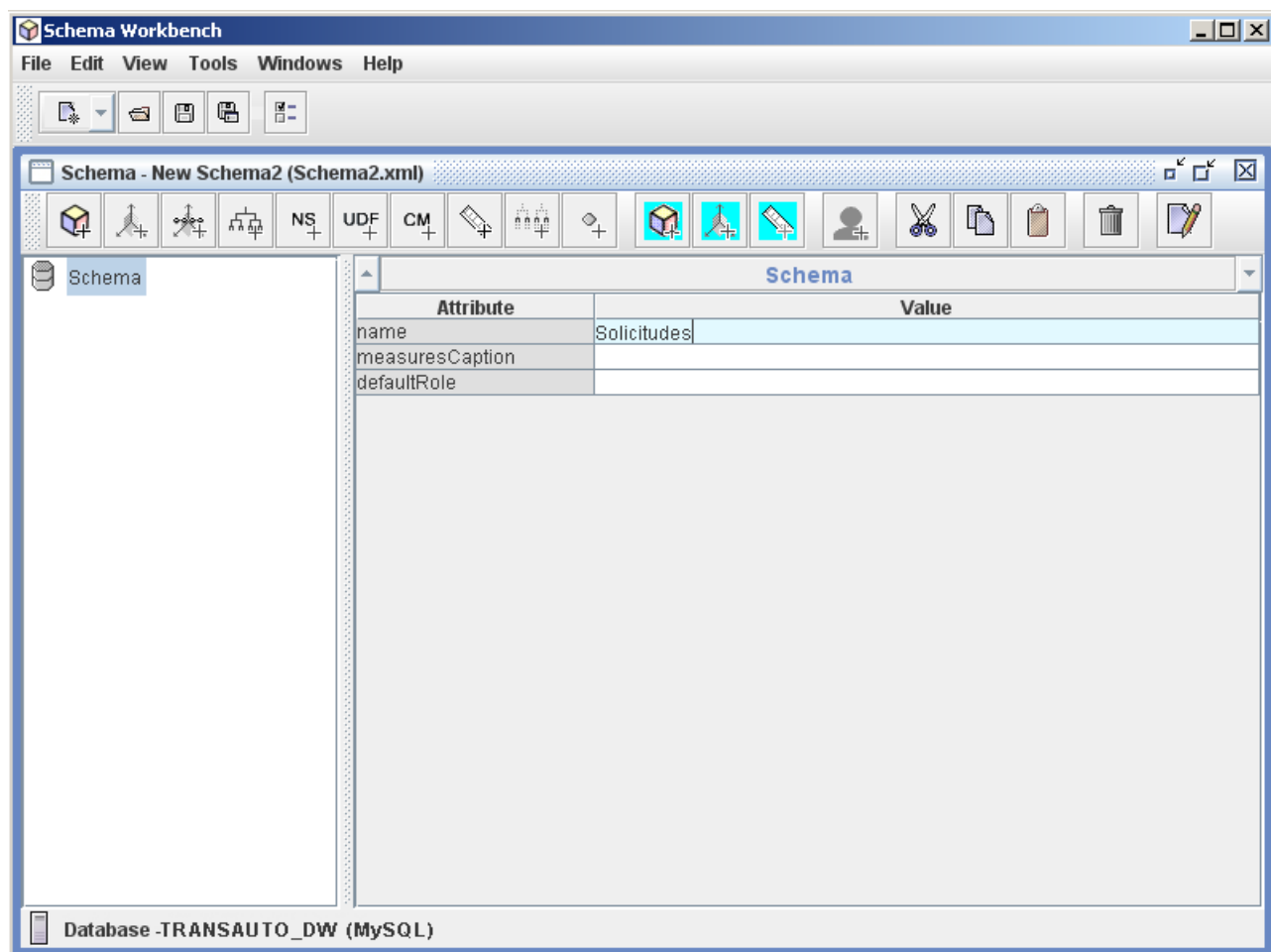


Queremos hacer un cubo que contenga la siguiente información:

Cantidad de solicitudes entradas al sistema, agrupadas por el tipo de solicitud y la fecha en que entró la solicitud, dicha fecha debe poseer la siguiente jerarquía:

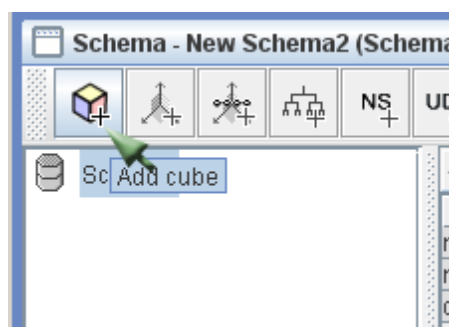
- 1- Año
 - 1.1- Mes
 - 1.1.1 Fecha

Empecemos entonces a ver como se usa esta herramienta para dar respuesta al cubo solicitado.

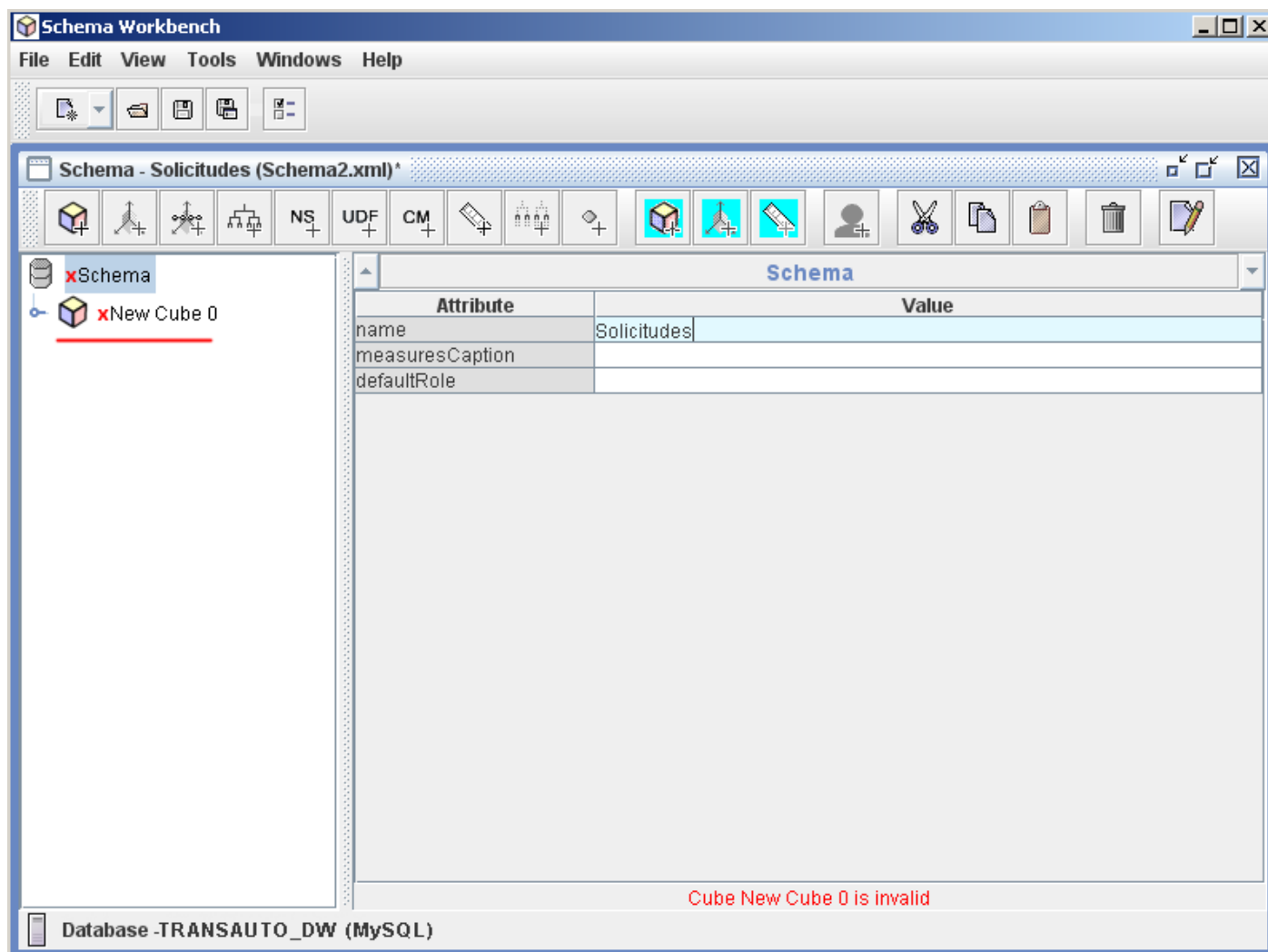


Como se puede apreciar se ha creado un “Schema” al cual se le pone un nombre en el campo “name”

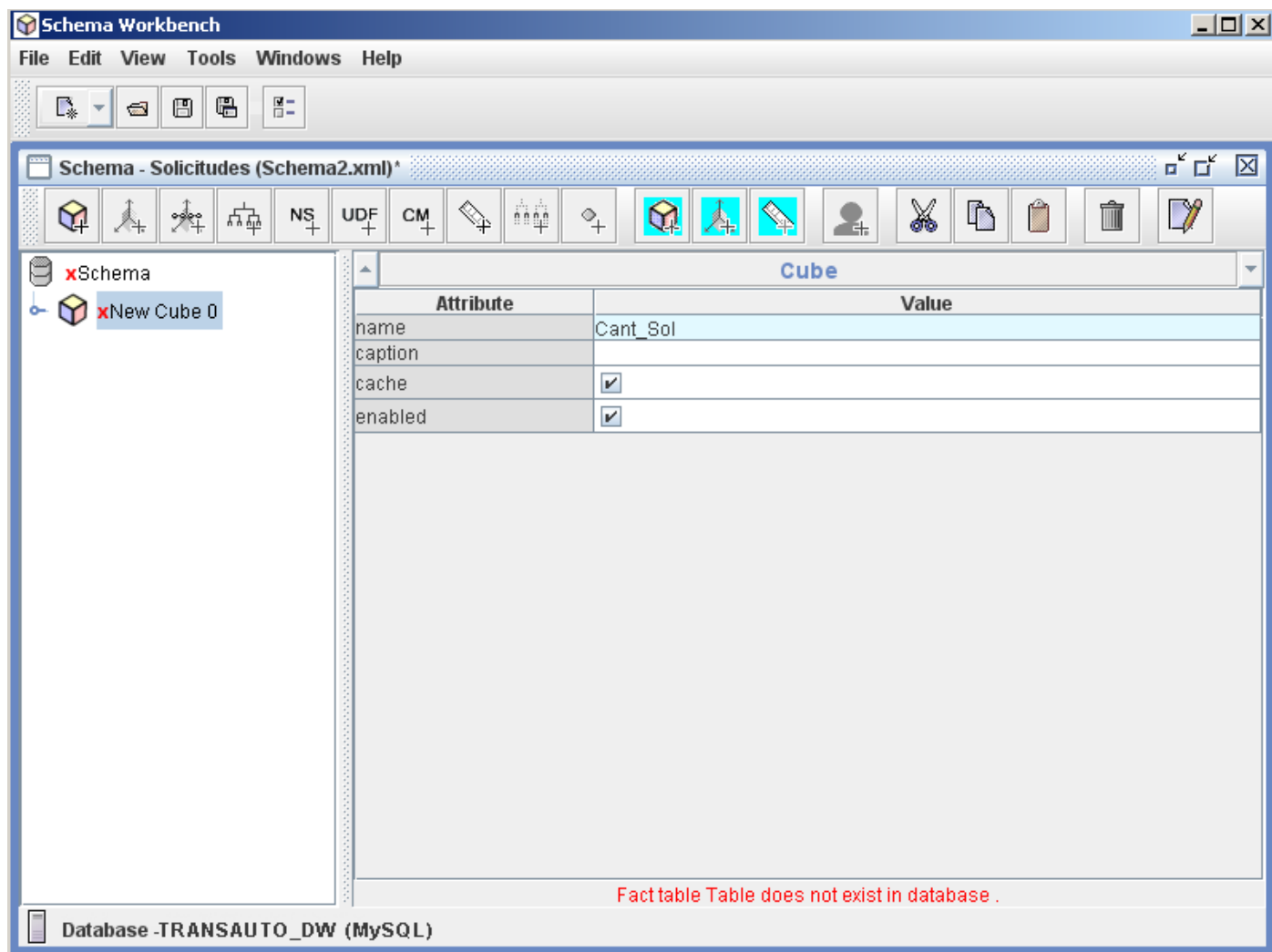
6) Nuevo Cubo:



Posteriormente con el schema marcado se le da al botón que apunta el cursor (Add Cube) y se crea un cubo.



A dicho cubo se le pone también un nombre en “name”

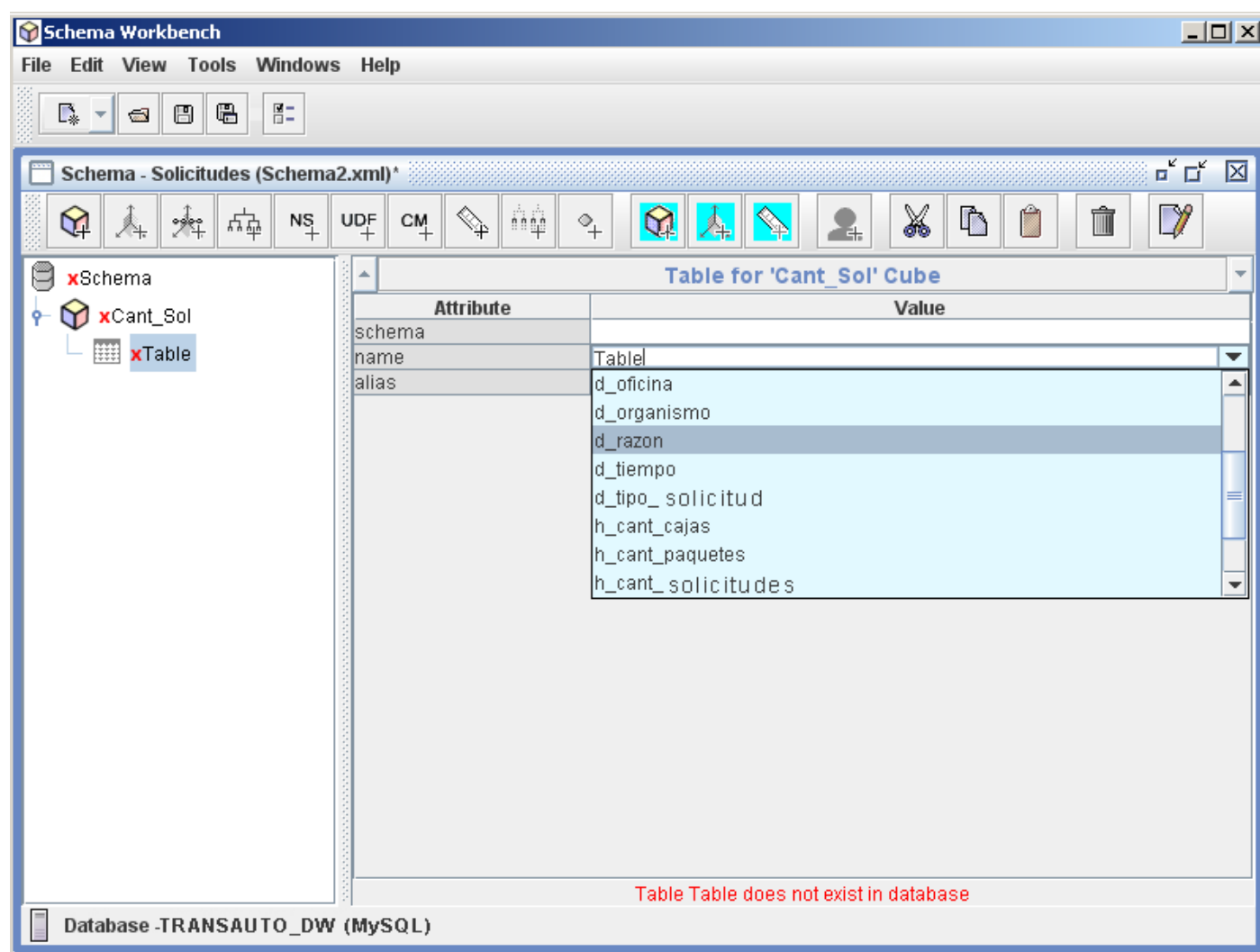


Como pueden apreciar el nombre dado a este cubo es “Cant_Sol”, es importante que observen los mensajes de color rojo que se muestra en el borde inferior de la ventana, en este caso “**Fact Table does not exist in database**” este mensaje hace referencia a que la tabla de hecho especificada no se encuentra en la base de datos, en realidad no hemos seleccionado ninguna, y hasta que no se haga se mantendrá el mensaje.

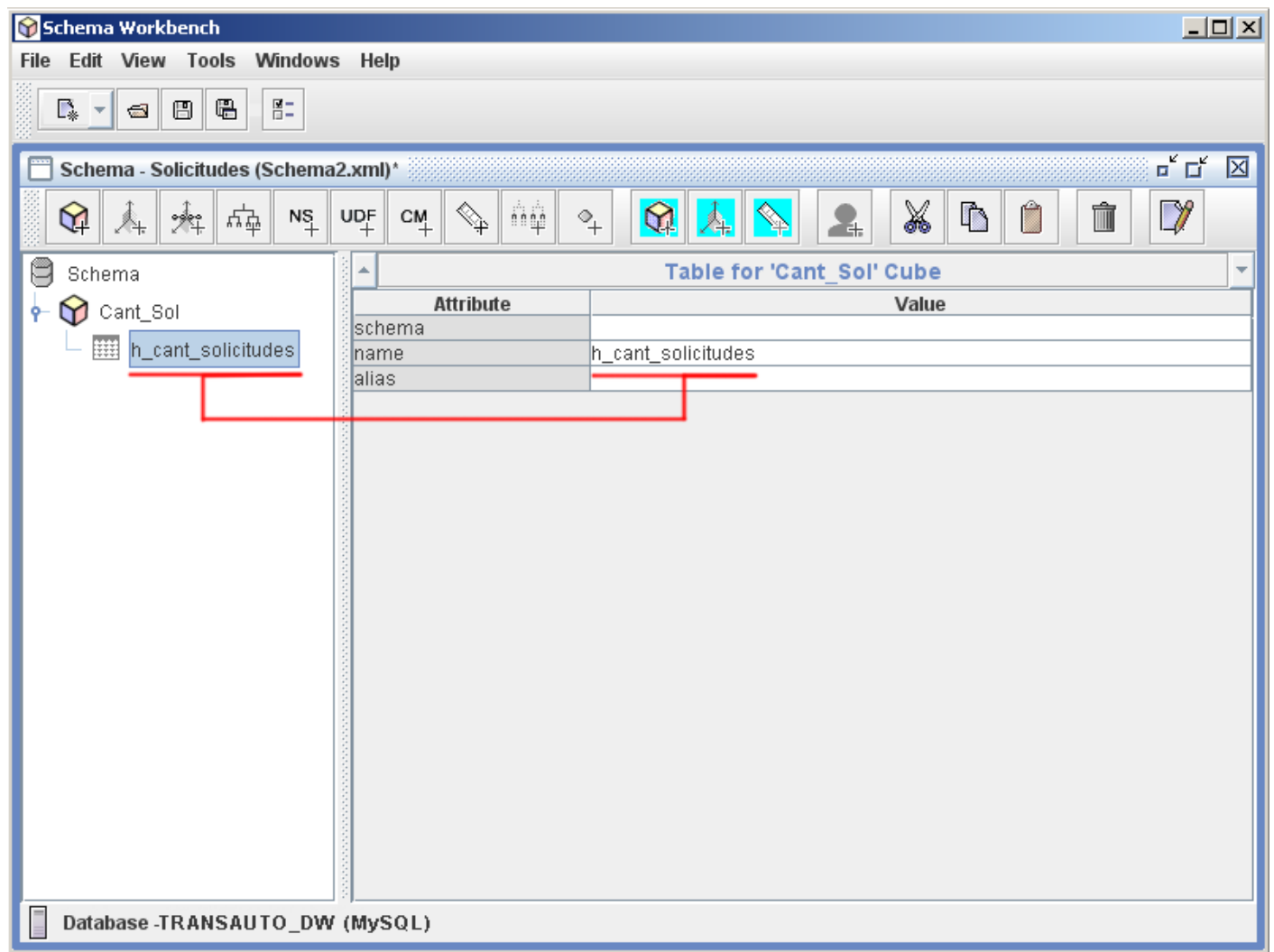
Siempre que aparezcan estos mensajes es que falta algún dato, por lo que el cubo está defectuoso.

7) Definir Tabla de Hechos:

Dando clic en el cubo “Cant_Sol” se puede ver como se despliega un icono de tabla “Table”, pues en este lugar se selecciona la tabla que contendrá los hechos, es decir aquello que deseamos analizar.

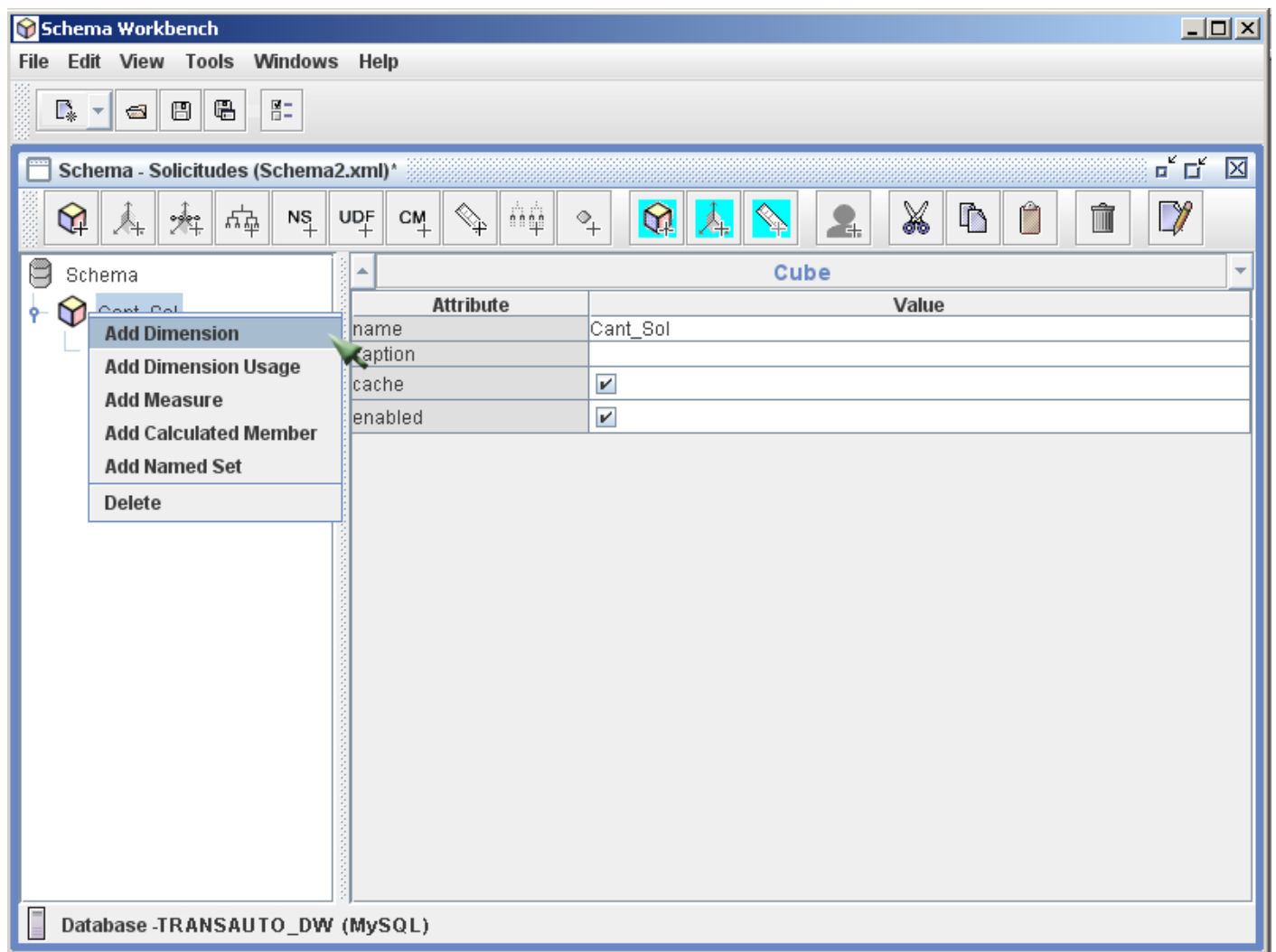


En esta imagen se puede apreciar como una vez seleccionado el nodo “Table” a la izquierda, se muestran a la derecha un conjunto de campos, entre los cuales está “name” que es donde se seleccionará la tabla de Hecho que se utilizará (solo se puede seleccionar alguna tabla que se encuentre en el listado mostrado).

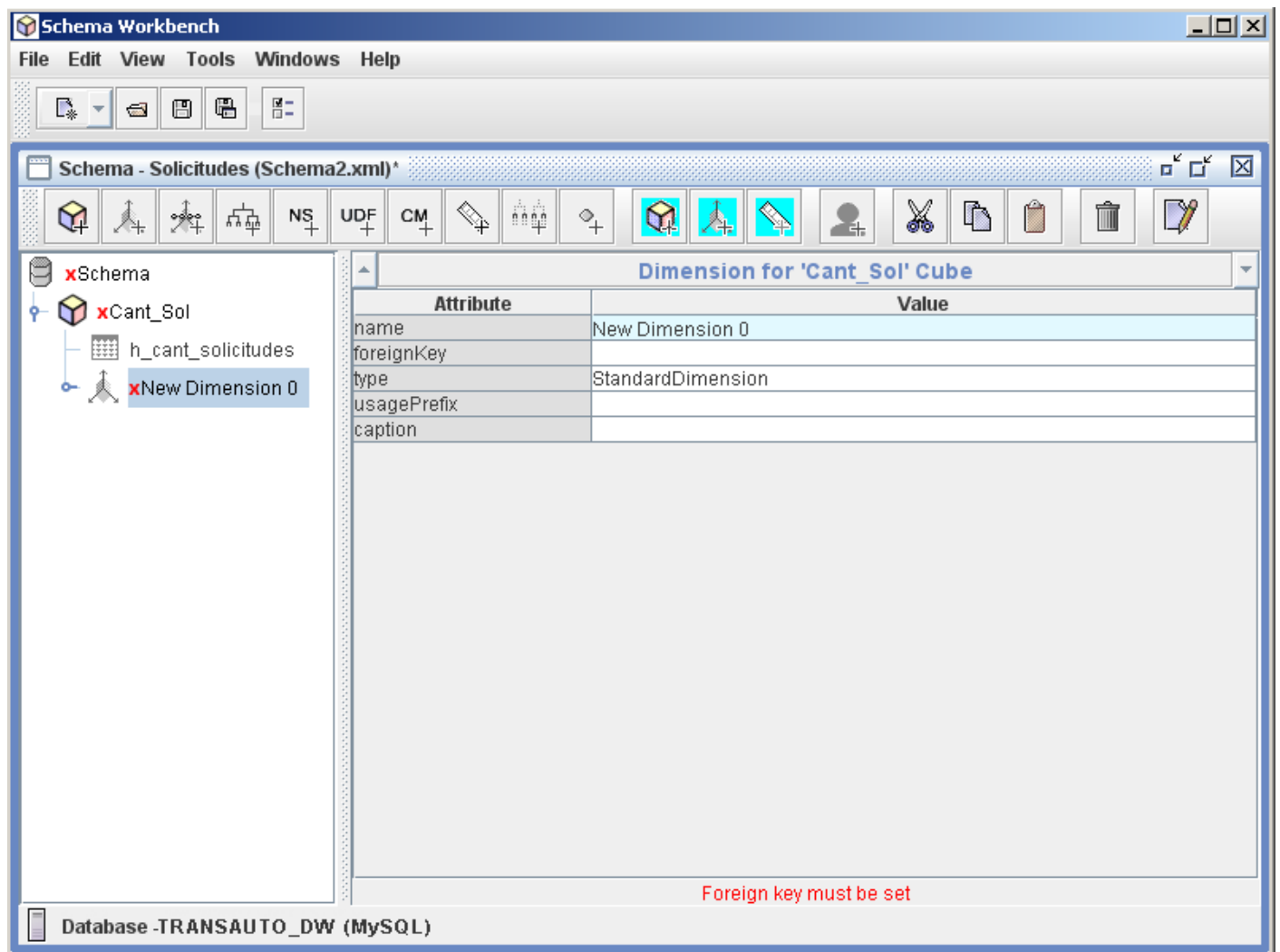


Vean como al seleccionar la tabla deseada “h_cant_solicitudes” el nodo “Table” se acoge al nombre de dicha tabla.

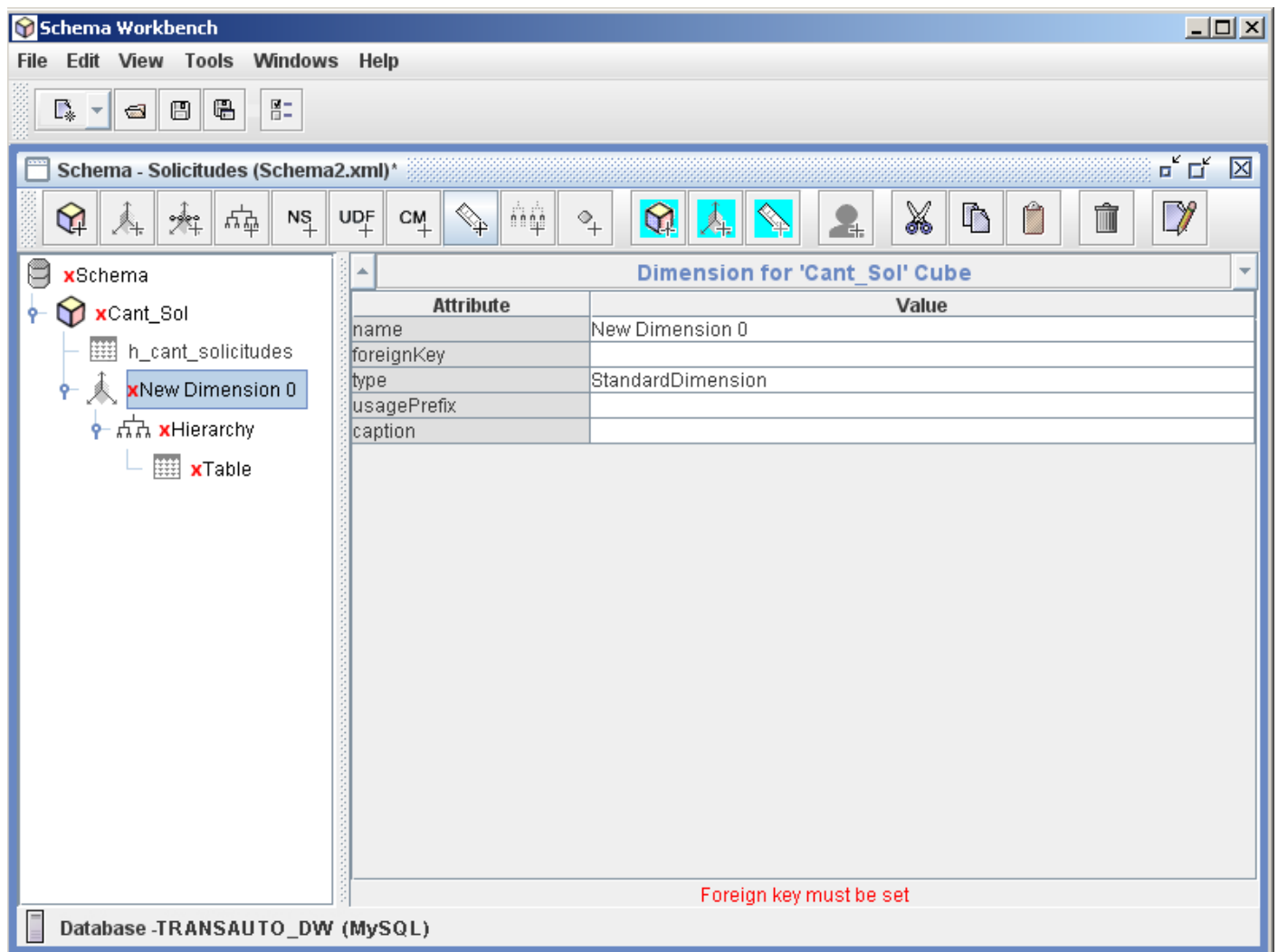
8) Añadir Dimensión “Tipos”:



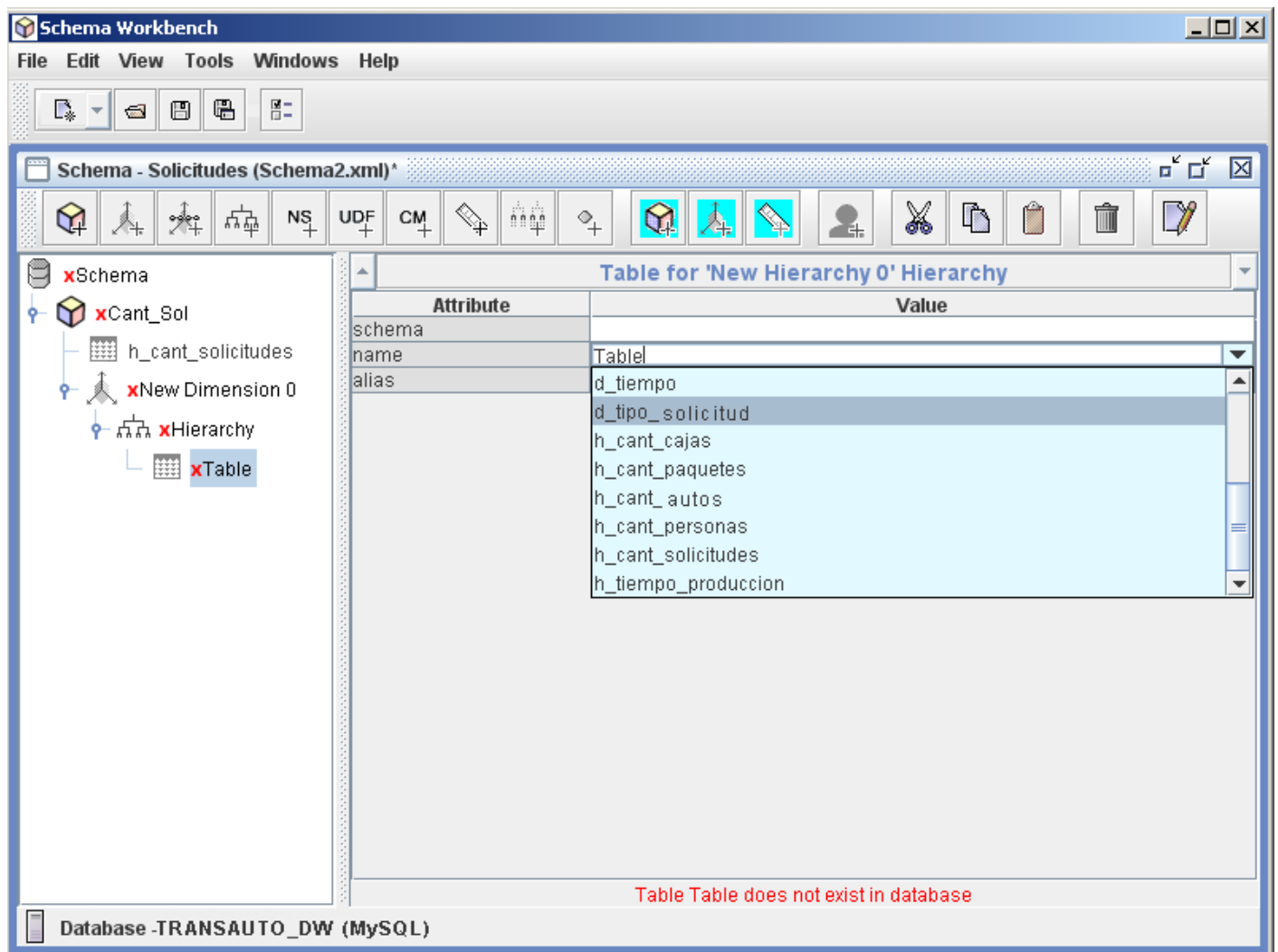
Acto seguido nos paramos encima del cubo otra vez, le damos clic derecho y del conjunto de opciones que sale se selecciona la opción “Add Dimension”.



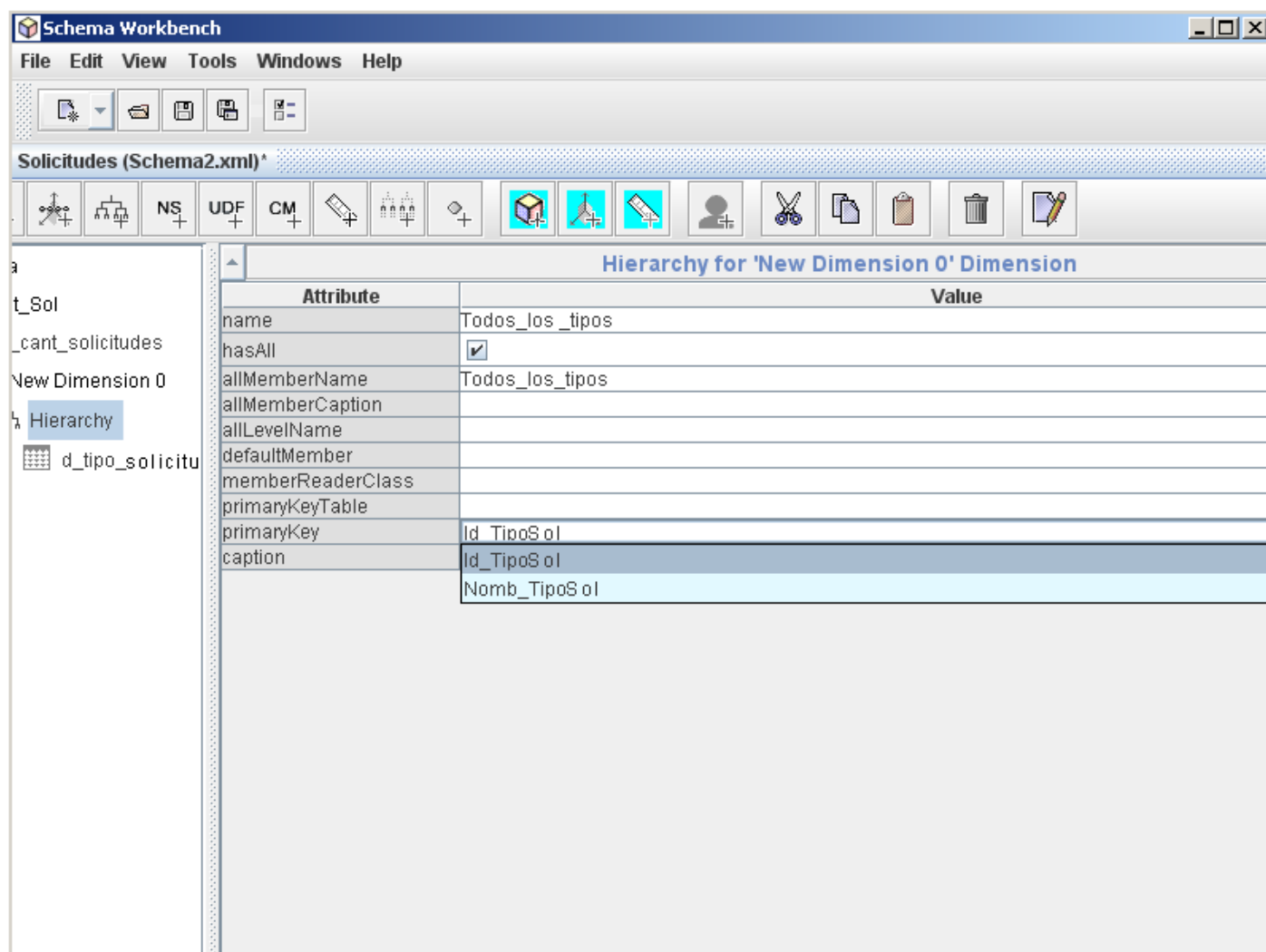
Como pueden apreciar ha surgido un nuevo elemento “New Dimensión 0”, es aquí donde se hará referencia a las tablas de “Dimensiones” que se quieran utilizar. Este elemento también tiene en su interior otras estructuras.



Aquí se muestra como el elemento "New Dimensión 0" tiene en su interior a "Hierarchy" (jerarquía), quien a su vez contiene un "Table", siempre me ha sido útil la filosofía de ir configurando estas estructuras compuestas, desde adentro hacia afuera, por lo que el modo de configuración es el siguiente.

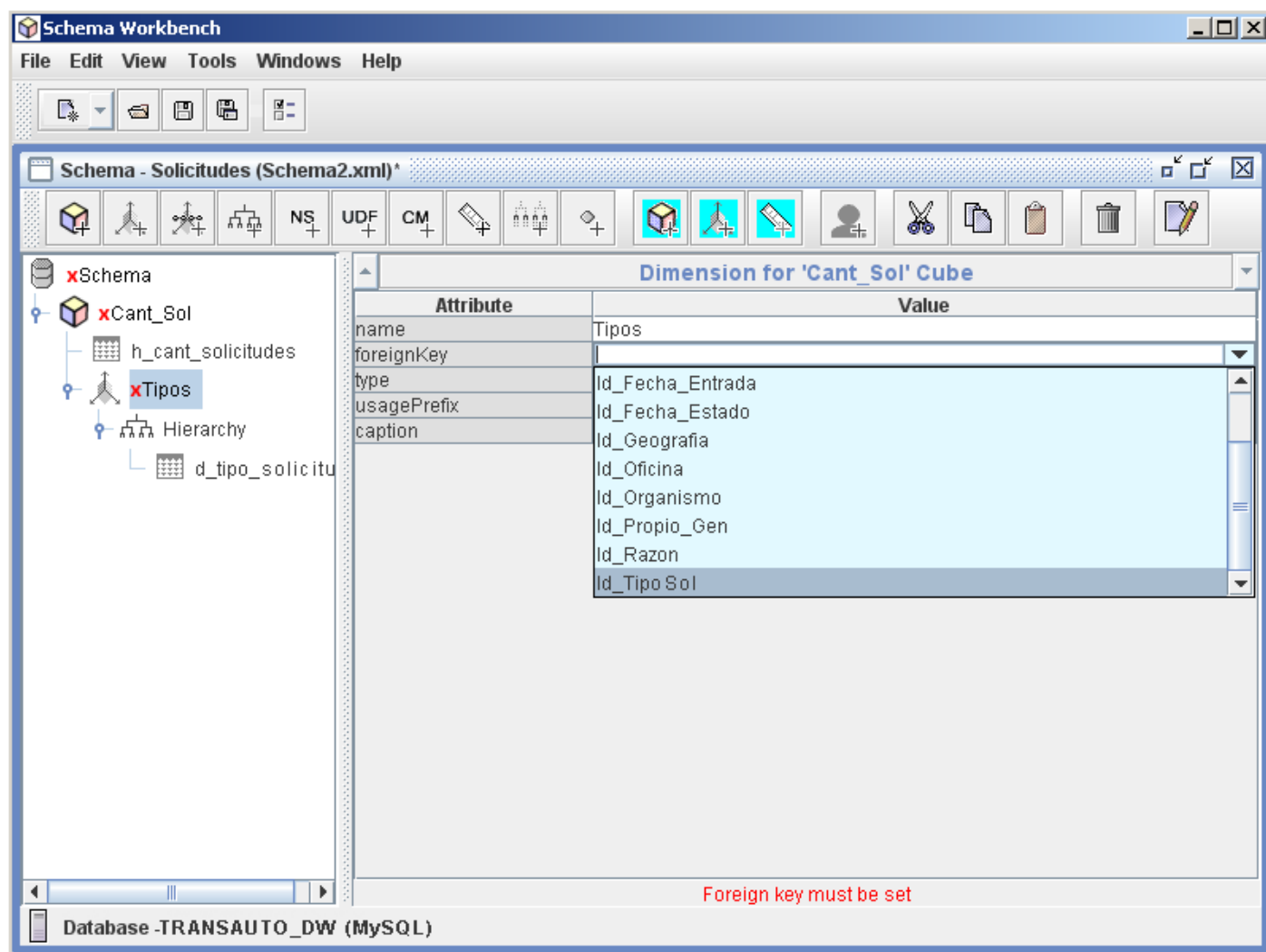


“Table” contiene el campo “name” en donde se escoge la tabla dimensión requerida, nos interesa el tipo de solicitud y el tiempo, así que seleccionamos la primera mencionada. De la misma forma que arriba “Table” se acoge al nombre de la tabla seleccionada.



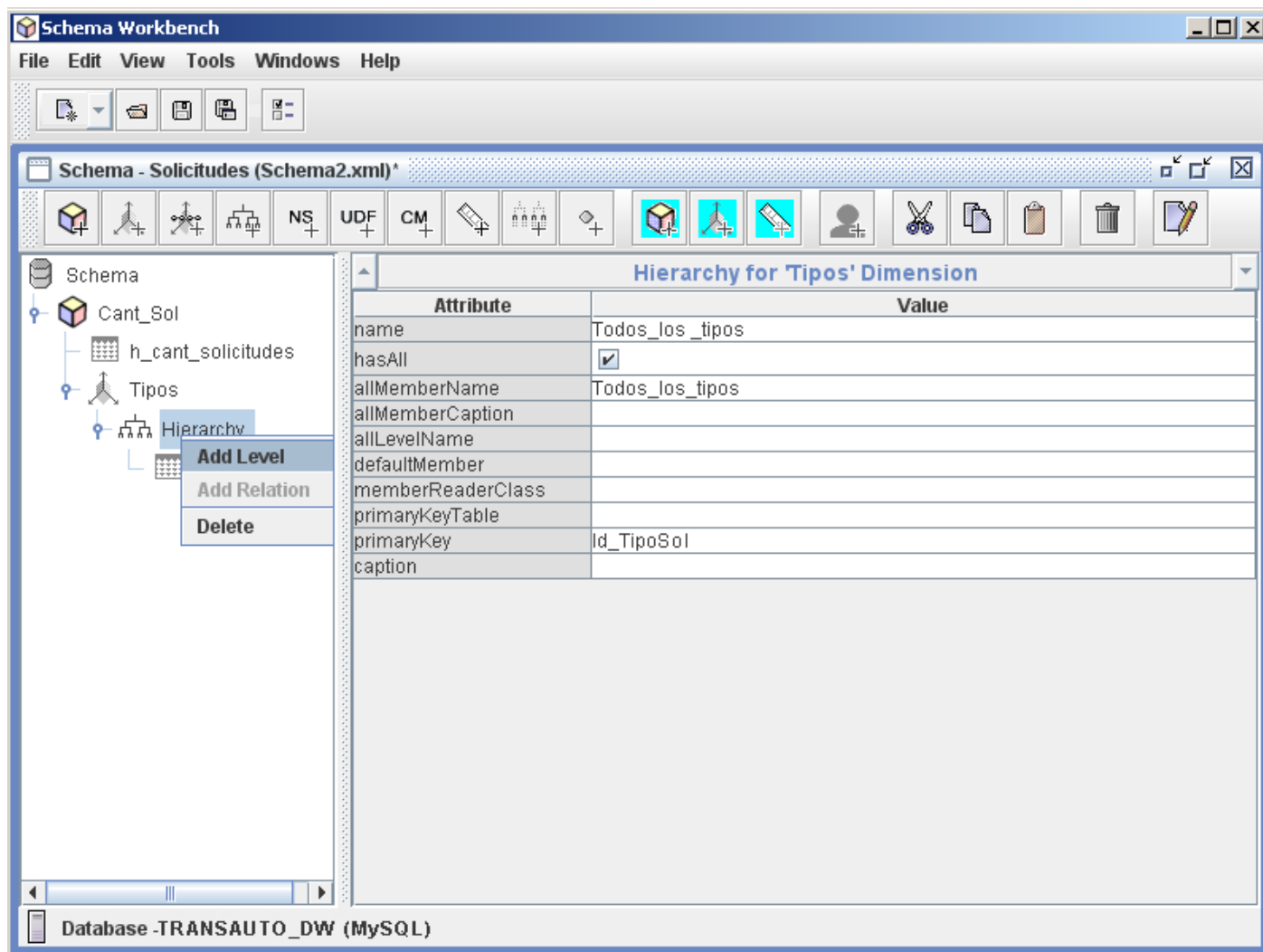
Luego nos paramos en “Hierarchy” y en los campos “name” y “allMemberName” ponemos “Todos_los_tipos” este será el título de los valores que saldrán en los reportes, vistas de análisis, etc.

En el campo “primaryKey” seleccionamos de los campos que tenga esta tabla dimensión aquel que sea la llave primaria, en este caso “Id_TipoSol”.

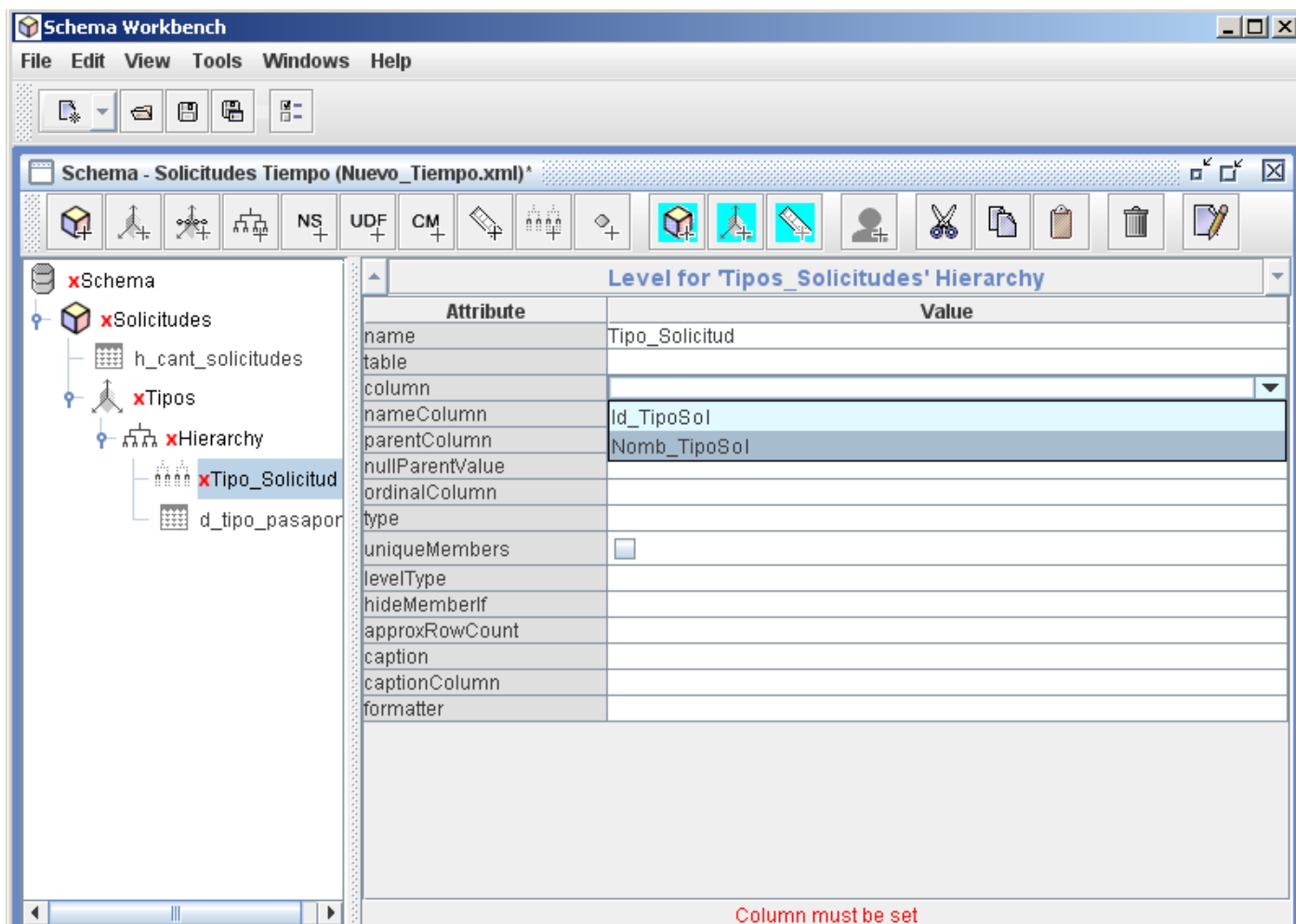


Seguidamente nos paramos sobre “New Dimensión 0” y le ponemos un nombre a este elemento en “name”.

Observen detenidamente ahora, se cuenta con un campo “foreingKey” eso se debe a que esta capa más arriba, está conectada a la tabla del cubo “h_cant_solicitudes” por lo que cuando vemos la lista que brinda “foreingKey” vemos todos los campos con los que cuenta la tabla hecho de este cubo, y como todos y todas sabemos las tablas hechos deben tener como llaves principales las llaves de las dimensiones que usa. De esta forma ya hemos establecido los campos por los que se indexan las tablas “h_cant_solicitudes” y “d_tipo_solicitud”.



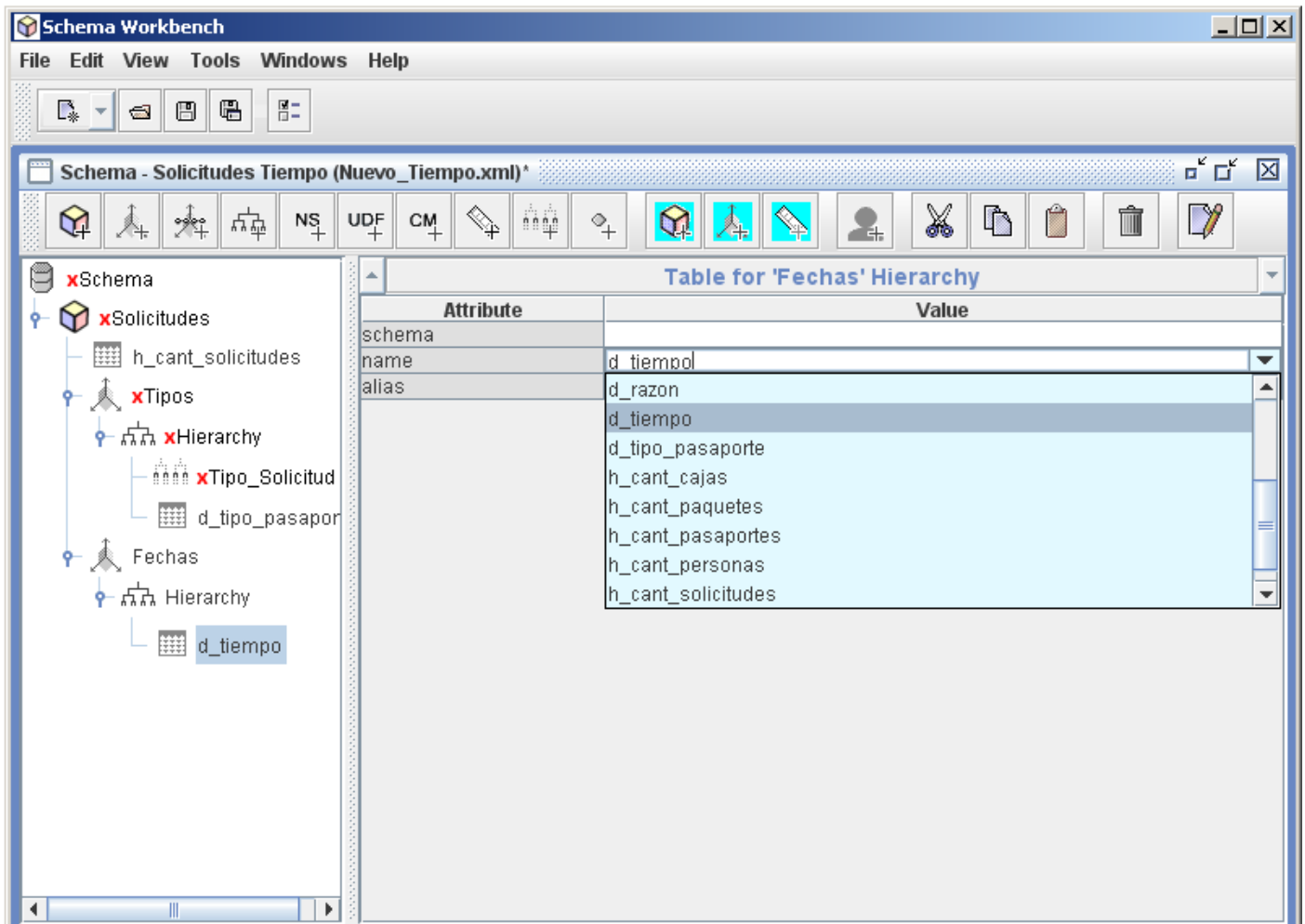
Bien, ya establecimos la conexión entre las tablas mencionadas, pero no hemos establecido que campo(s) son los que mostraremos. Para ellos nos paramos en “Hierarchy” y damos clic derecho y seleccionamos “Add Level”.



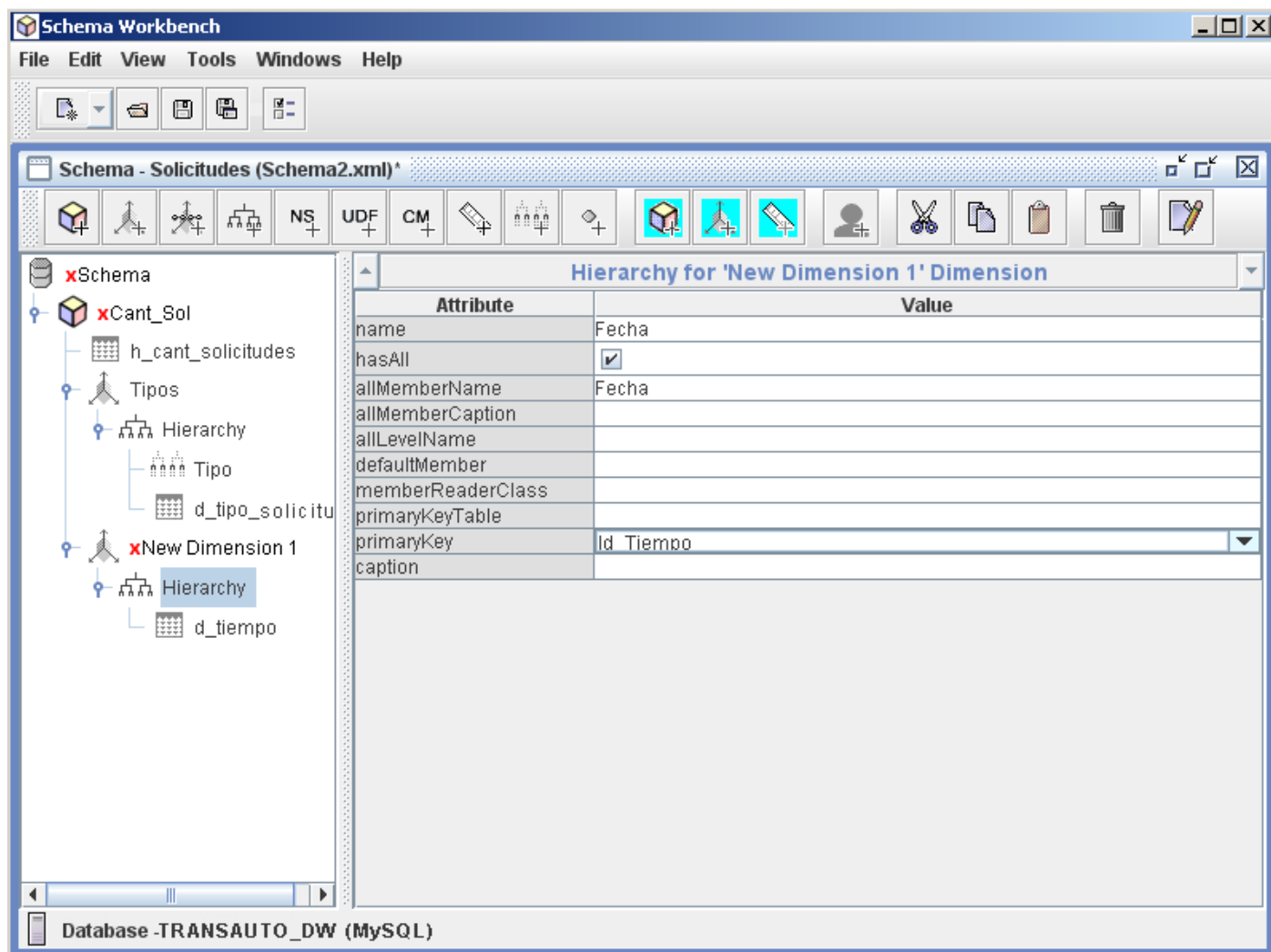
A este “Level” le ponemos nombre en “name” y en “nameColumn” seleccionamos el campo que deseamos se muestre.

9) Añadir Dimensión “Fechas”:

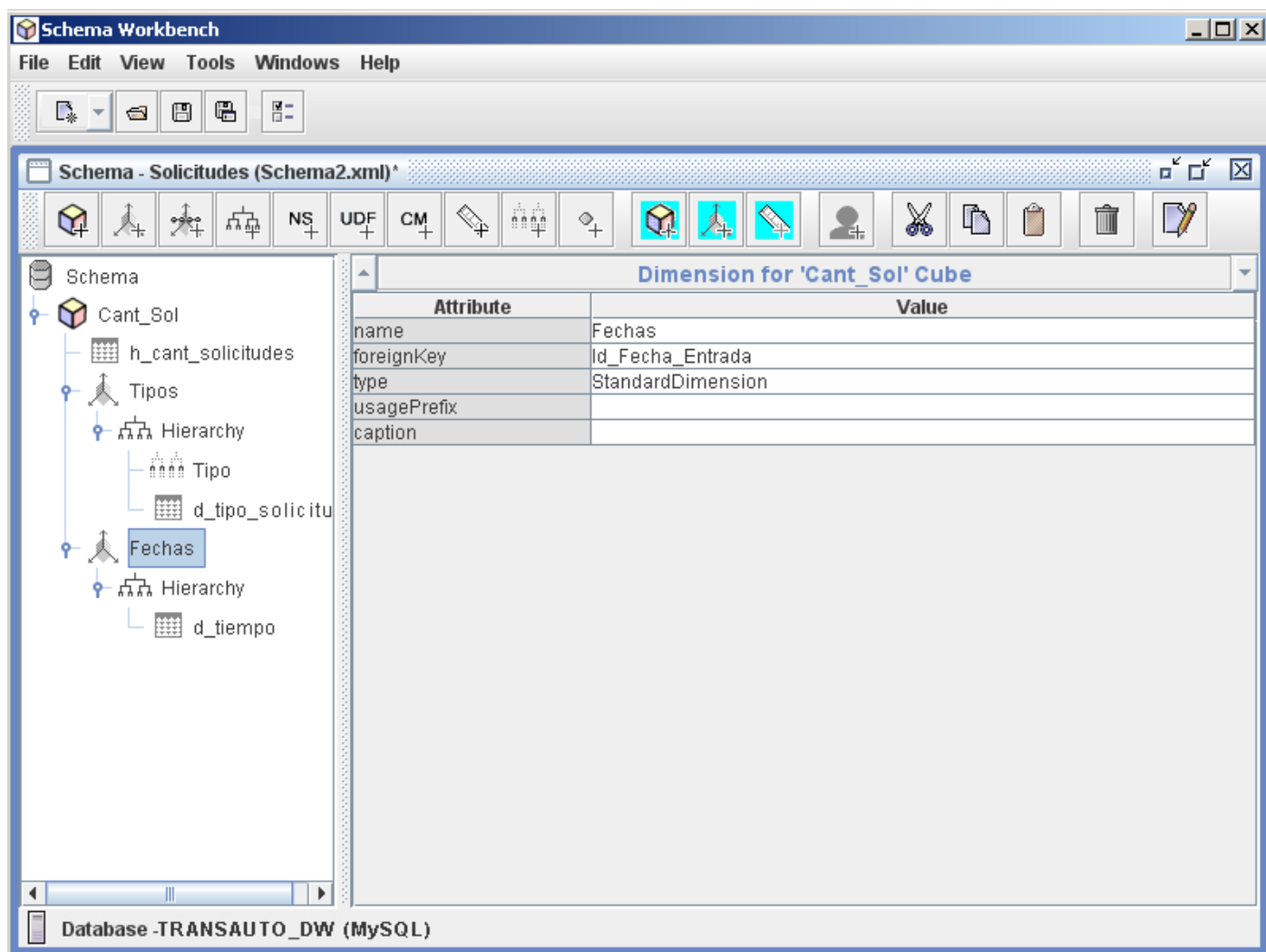
Ya con esto hemos terminado de configurar la dimensión del tipo de solicitud, según lo requerido nos queda crear otra para el tiempo, para lo cual nos volvemos a posicionar arriba del cubo, le damos clic derecho, creamos una nueva dimensión (“Add Dimension”) y bajamos hasta el nivel más bajo “Table” y seleccionamos la tabla correspondiente.



Se selecciona la tabla “d_tiempo”.

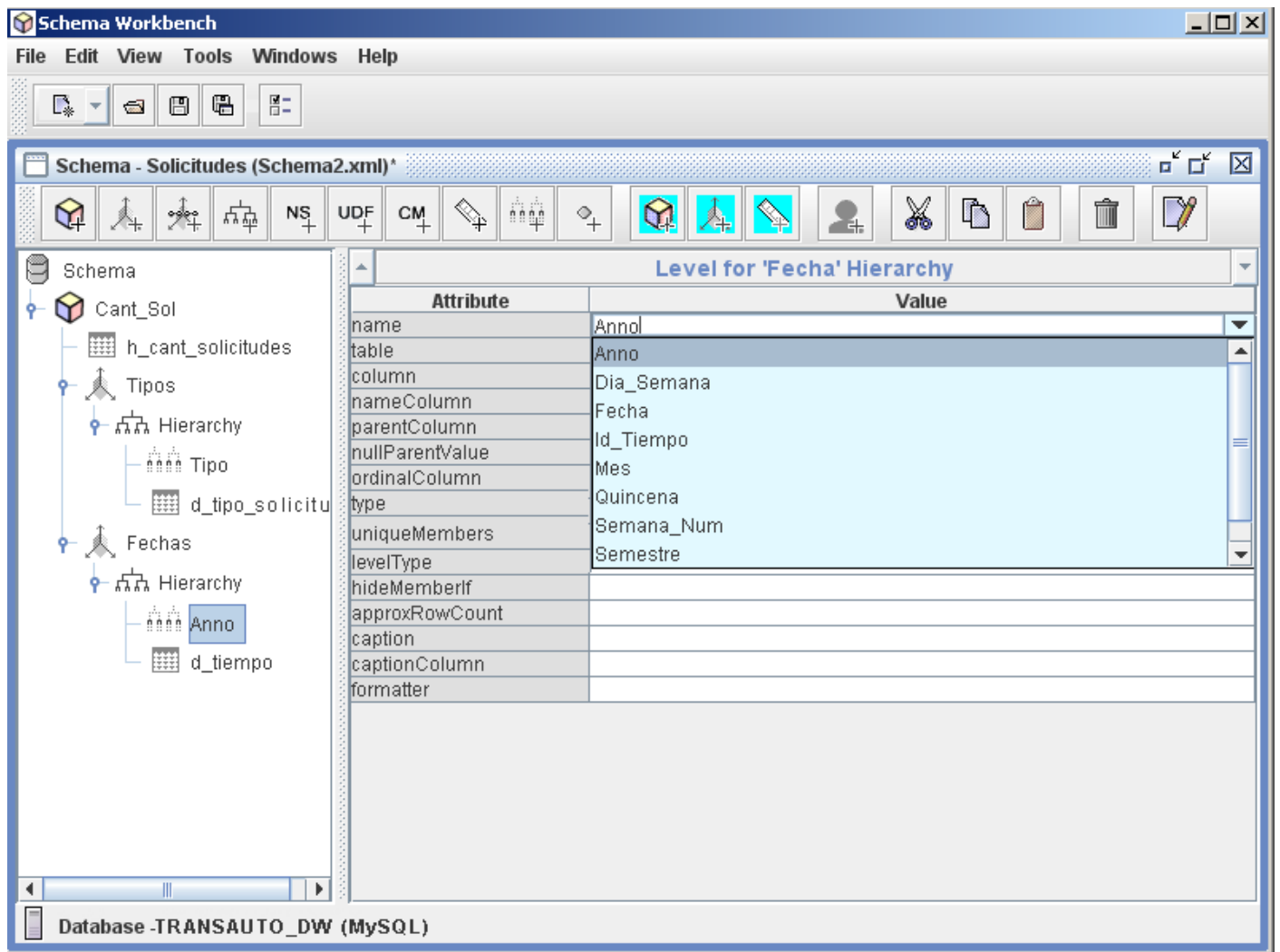


Configuramos la jerarquía "Hierarchy" como se muestra arriba en la foto.

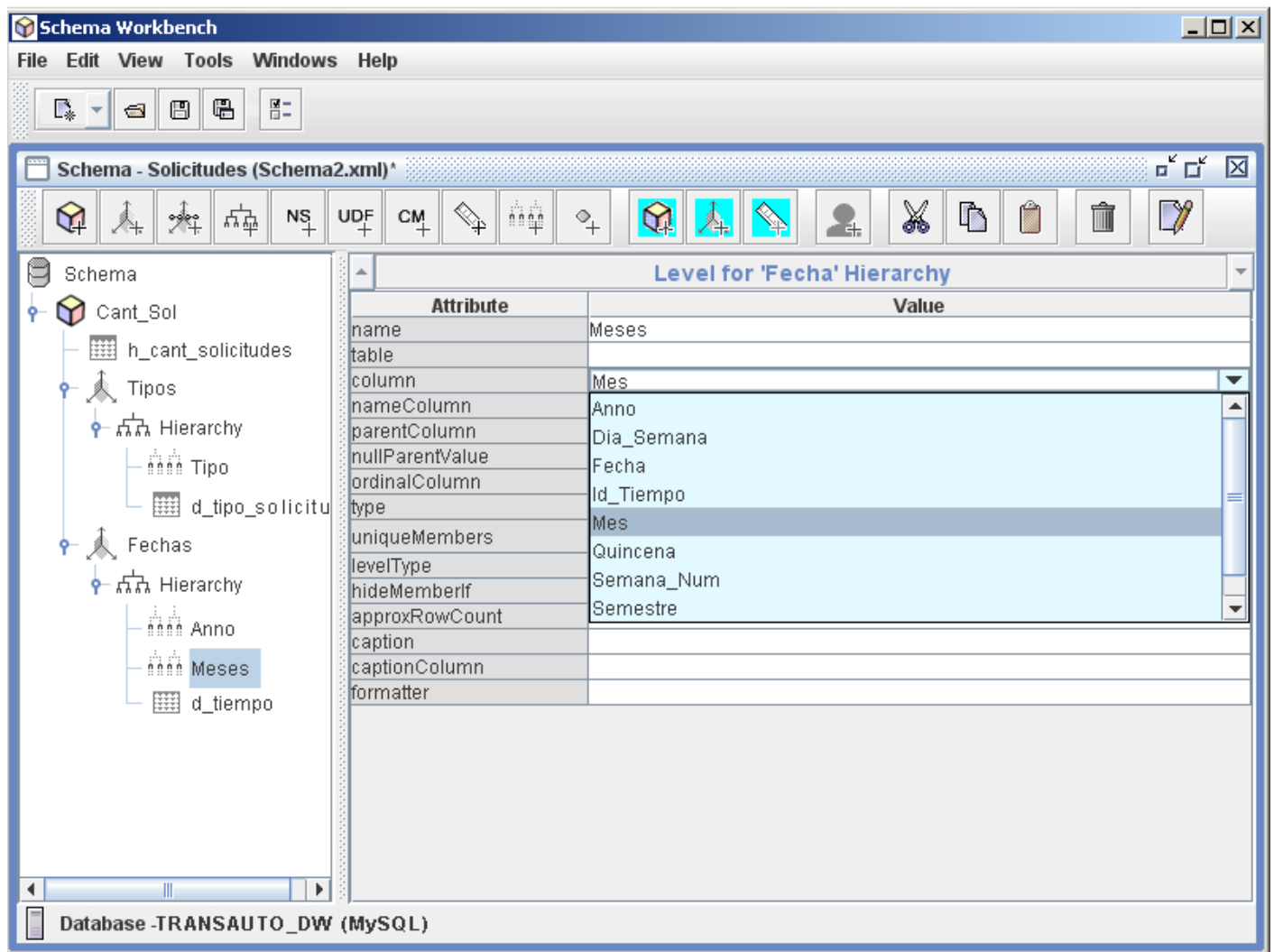


Y la “New Dimension” como se muestra en la figura de arriba, fíjense que se ha seleccionado el campo “Id_Fecha_Entrada” pues es lo que se solicita.

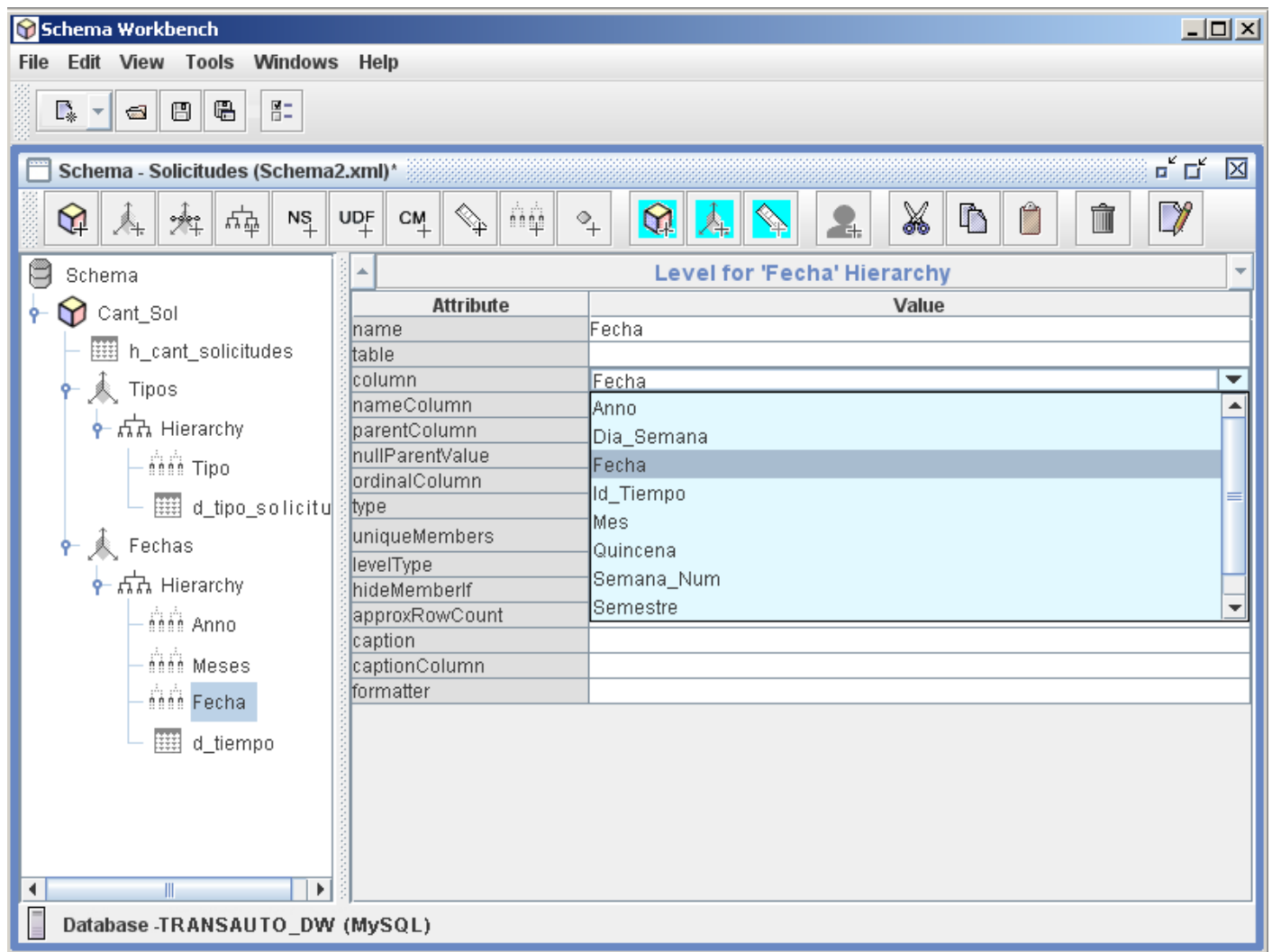
Ahora esta dimensión “Fechas” tiene una pequeña diferencia con respecto a “Tipos” y es que aquí se desean mostrar más de un nivel, pues se desea organizar por año, por cada año los meses, y por las fechas. A continuación veremos como se debe hacer para lograr esto.



Se crea un primer nivel que va a coger el campo “Anno”, que referencia al año.



Luego se crea otro nivel que guardará el campo referente a los meses: "Mes".

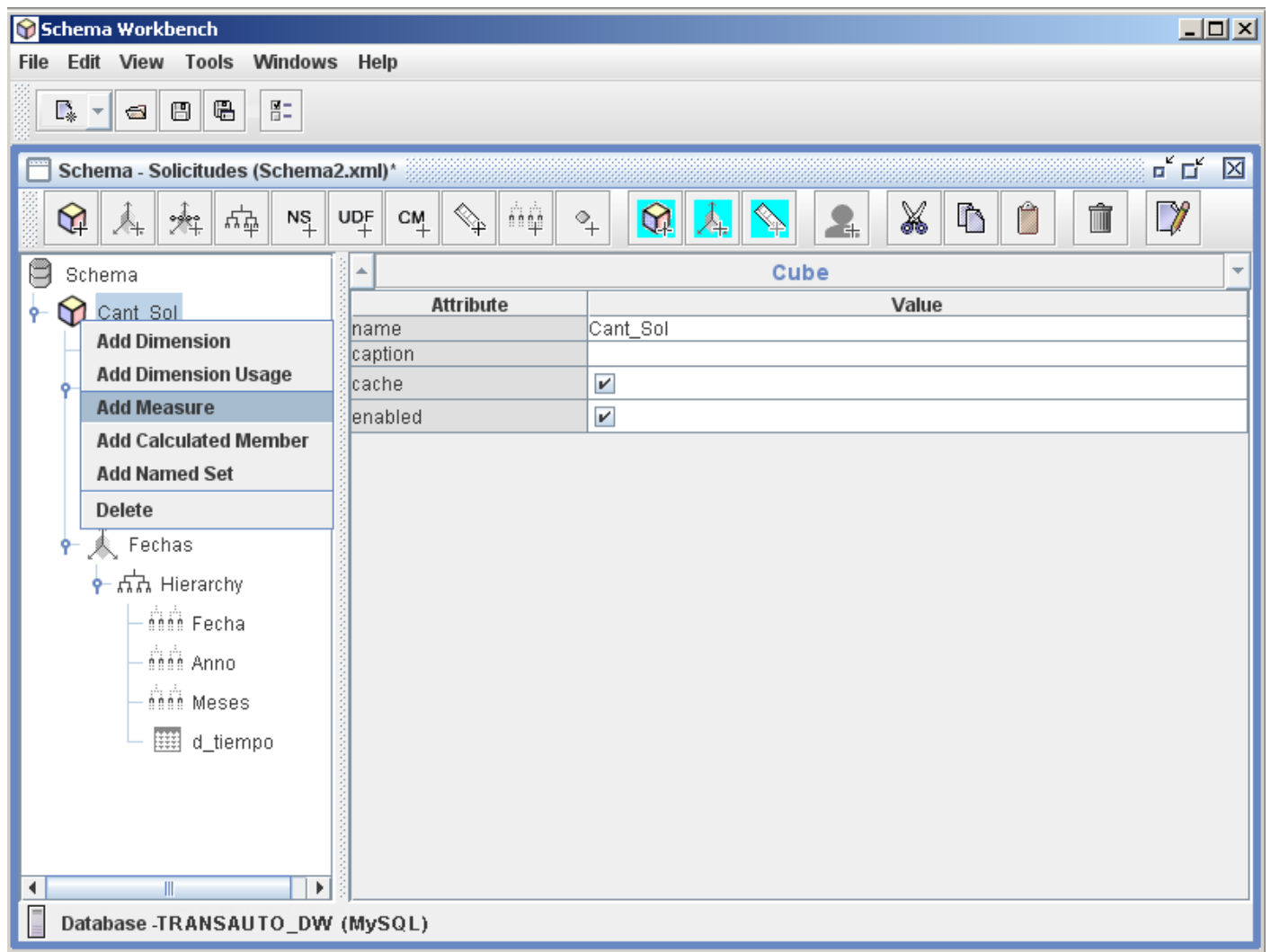


Y por último el nivel “Fecha” que acogerá el campo “Fecha”.

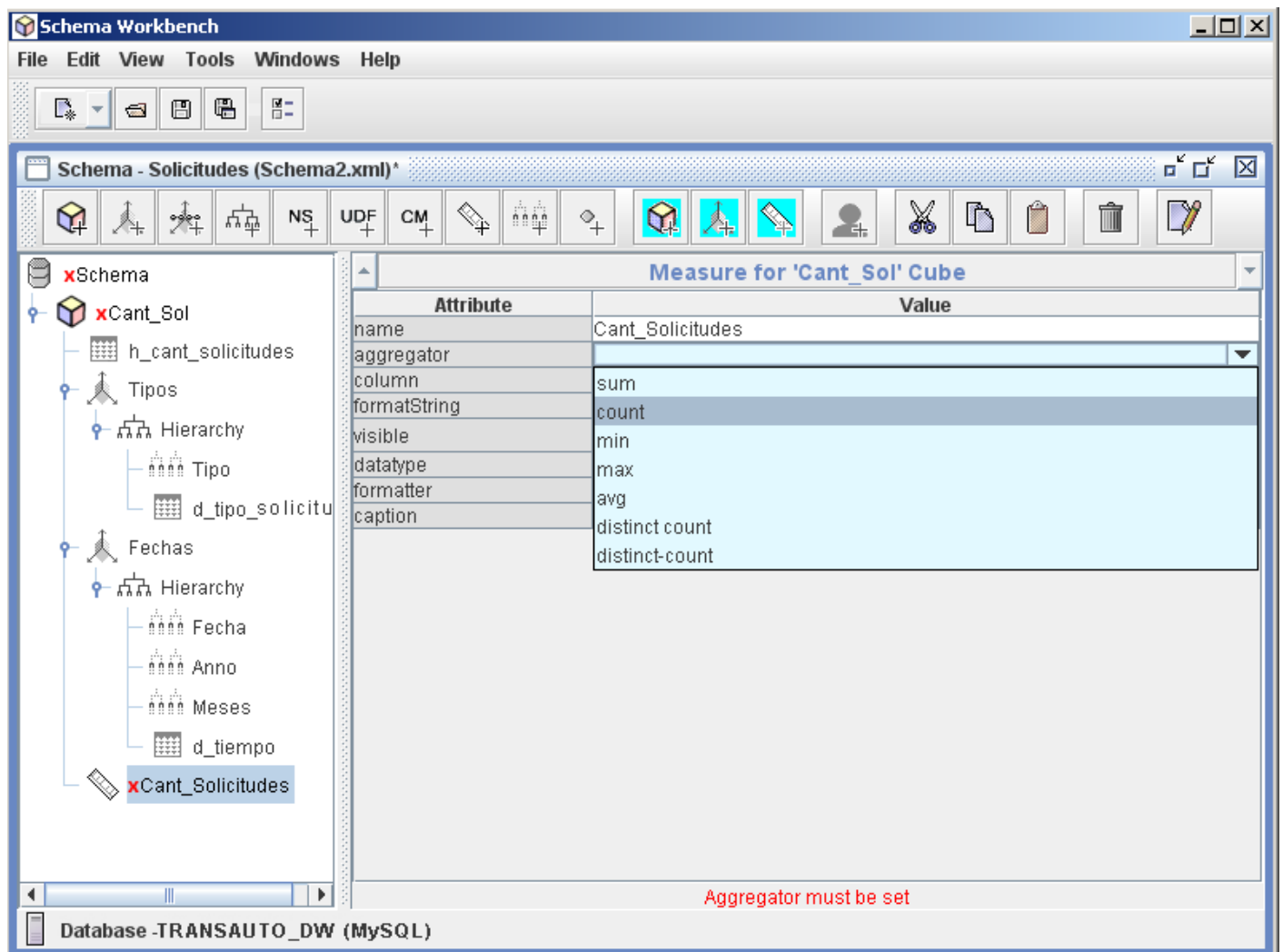
10) Añadir Indicador:

Bien, ya tenemos modelado nuestro cubo con las especificaciones requeridas, ahora solo nos queda hacer el cálculo de las solicitudes realizadas.

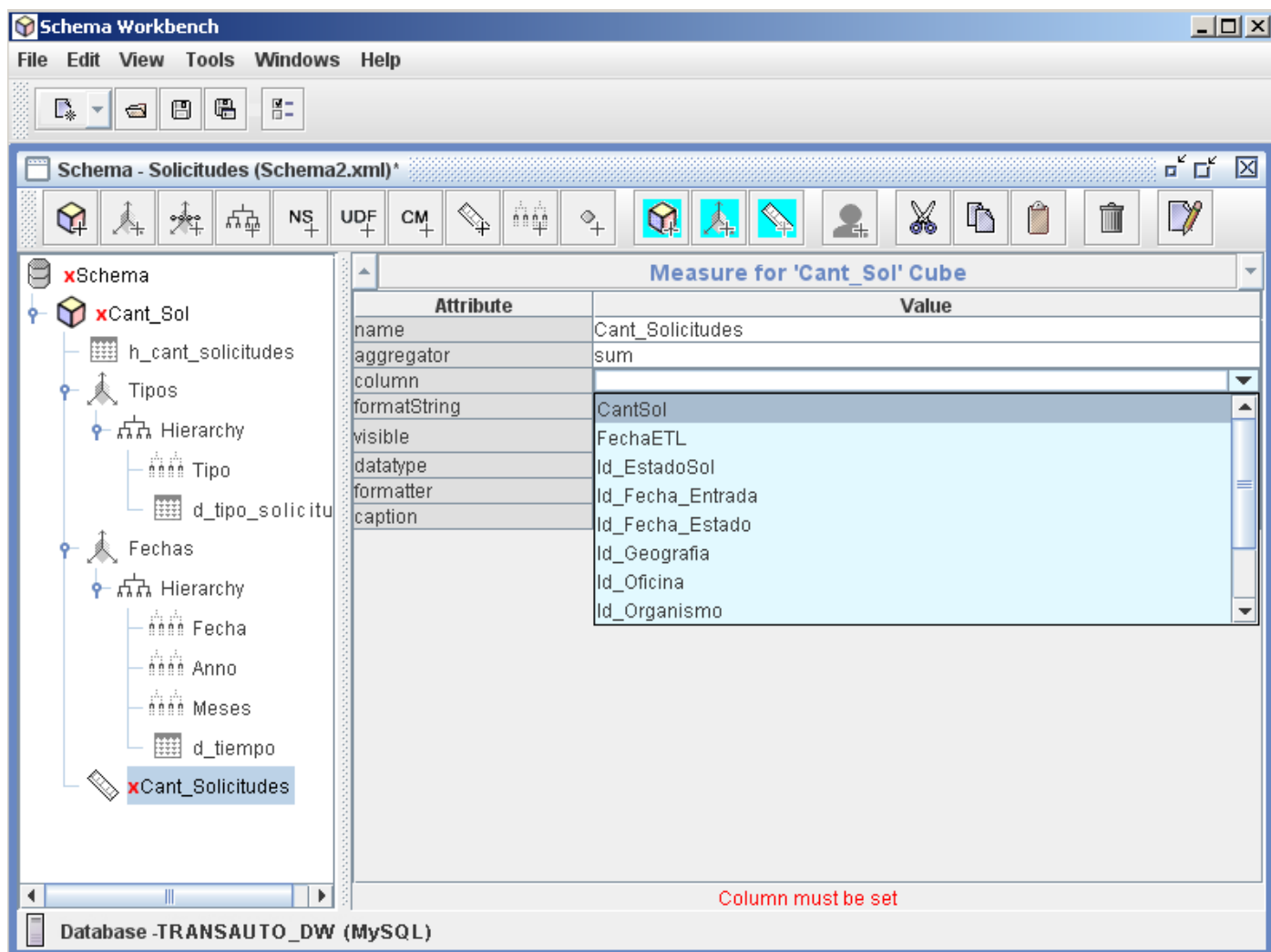
Esto se hace con el campo “Cant_sol” que tiene la tabla hecho “h_cant_solicitudes” de la siguiente forma.



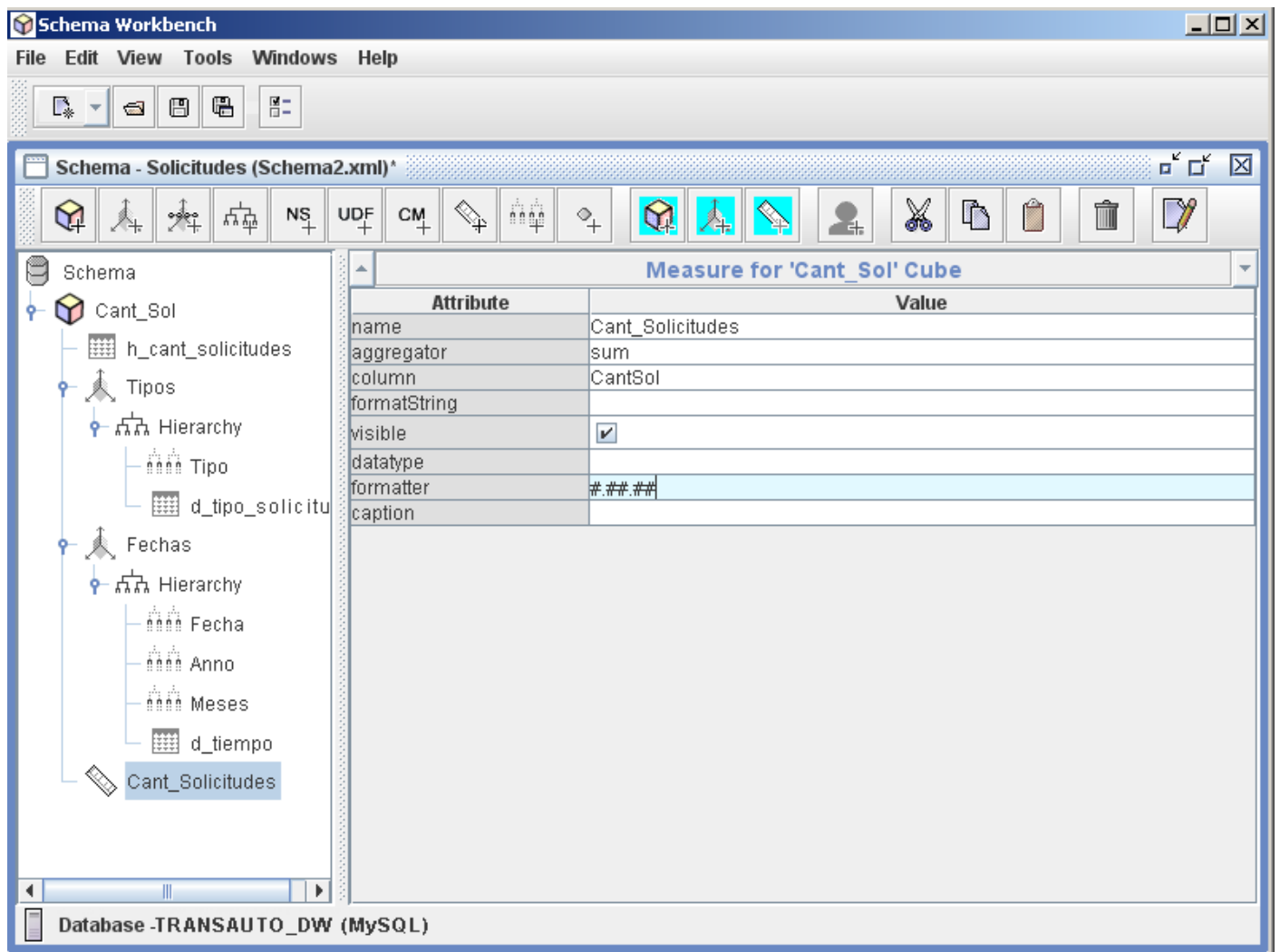
Nos posicionamos arriba del cubo, damos clic derecho y seleccionamos “Add Measure”.



En “name” ponemos el nombre de este valor a calcular, en “aggregator” seleccionamos la operación de sumariazación que se desee realizar, en nuestro caso “Sum” (sumar).

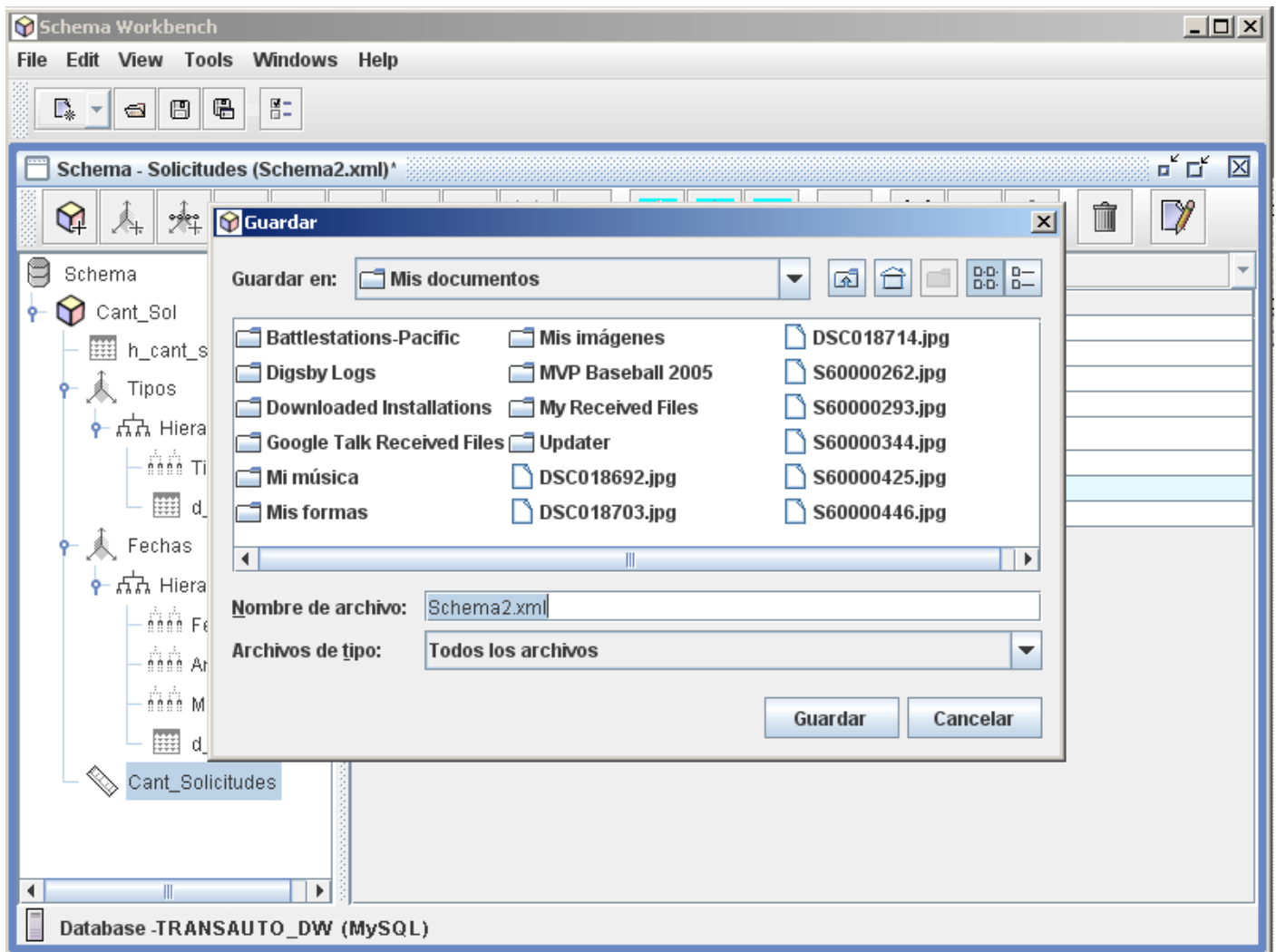


Luego, en “column” seleccionamos el campo “CantSol”, que será el hecho al que le vamos a realizar el cálculo seleccionado anteriormente.



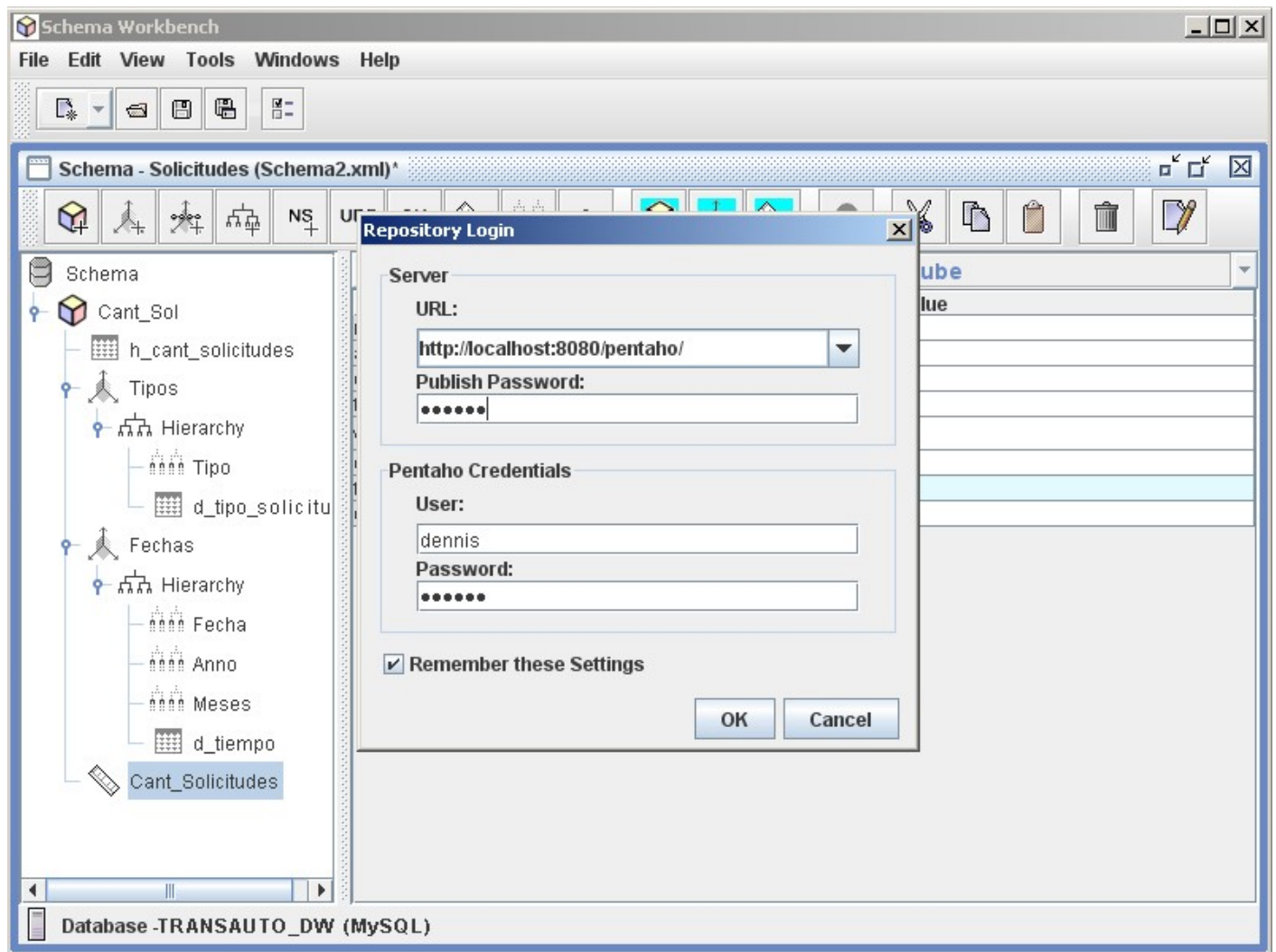
Aquí especificamos el formato que le vamos a dar al resultado calculado, entero, decimal, con símbolos como "\$, %" en dependencia de lo que se calcule.

11) Guardar el Cubo:



Solo nos queda salvar este cubo, el cual tendrá una extensión “.xml”.

12) Publicar el Cubo:



Y por último publicar el cubo en nuestro servidor Pentaho.

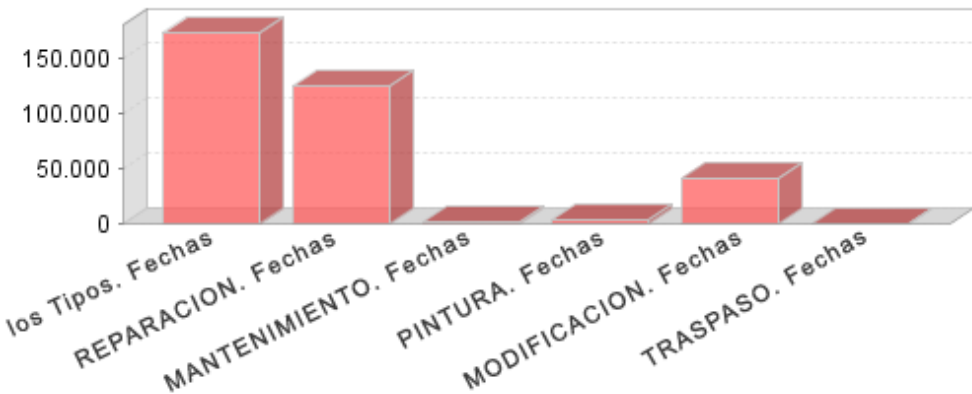
13) Ejemplos de Análisis Interactivos:

Con el cubo creado se podrán obtener vistas como las siguientes:

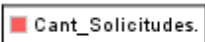
Analysis View

		Medidas
Tipo_Pasaporte	Tiempo	● Cant_Solicitudes
<input type="checkbox"/> Todos los Tipos	<input type="checkbox"/> Fechas	172399
REPARACION	<input type="checkbox"/> Fechas	124623
MANTENIMIENTO	<input type="checkbox"/> Fechas	1823
PINTURA	<input type="checkbox"/> Fechas	3788
MODIFICACION	<input type="checkbox"/> Fechas	41967
TRASPASO	<input type="checkbox"/> Fechas	198

Slicer:



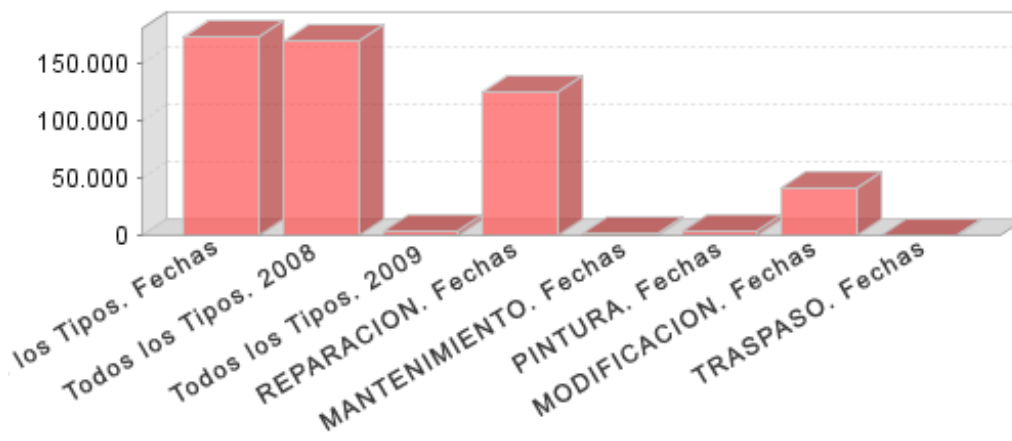
Slicer:



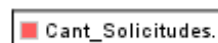


		Medidas
Tipo_Pasaporte	Tiempo	● Cant_Solicitudes
<input type="checkbox"/> Todos los Tipos	<input type="checkbox"/> Fechas	172399
	<input checked="" type="checkbox"/> 2008	168702
	<input checked="" type="checkbox"/> 2009	3697
REPARACION	<input checked="" type="checkbox"/> Fechas	124623
MANTENIMIENTO	<input checked="" type="checkbox"/> Fechas	1823
PINTURA	<input checked="" type="checkbox"/> Fechas	3788
MODIFICACION	<input checked="" type="checkbox"/> Fechas	41967
TRASPASO	<input checked="" type="checkbox"/> Fechas	198

Slicer:



Slicer:



Analysis View		
Tipo_Pasaporte	Tiempo	Medidas
<input type="checkbox"/> Todos los Tipos	<input type="checkbox"/> Fechas	● Cant_Solicitudes
	<input type="checkbox"/> Fechas	172399
	<input type="checkbox"/> 2008	168702
	<input type="checkbox"/> 2009	3697
	<input type="checkbox"/> Abril	283
	<input type="checkbox"/> Agosto	1
	<input type="checkbox"/> Enero	1037
	<input type="checkbox"/> Febrero	1169
	<input type="checkbox"/> Junio	5
	<input type="checkbox"/> Marzo	1202
REPARACION	<input type="checkbox"/> Fechas	124623
MANTENIMIENTO	<input type="checkbox"/> Fechas	1823
PINTURA	<input type="checkbox"/> Fechas	3788
MODIFICACION	<input type="checkbox"/> Fechas	41967
TRASPASO	<input type="checkbox"/> Fechas	198

Slicer:

