
Práctica-5: Herramientas OLAP Mondrian.



**UNIVERSIDAD
DE GRANADA**

Sistemas Multidimensionales (2019-2020)

Daniel Bolaños Martínez
danibolanos@correo.ugr.es
Grupo 2 - Viernes 15:30h

Índice

1. Define el esquema multidimensional para Mondrian asociado a la BD PostgreSQL creada en la actividad Herramientas ETL: PDI (Pentaho Data Integration).	3
2. Sobre Mondrian, utilizando el esquema multidimensional:	24
2.1. Obtén un informe inicial libre y explica su contenido.	26
2.2. Mediante Drill-down obtén un nuevo informe y explica su contenido. . .	27
2.3. Mediante Slice&Dice obtén un nuevo informe y explica su contenido. .	29
2.4. Mediante Roll-up obtén un nuevo informe y explica su contenido. . . .	30
3. Bibliografía.	31

1. Define el esquema multidimensional para Mondrian asociado a la BD PostgreSQL creada en la actividad Herramientas ETL: PDI (Pentaho Data Integration).

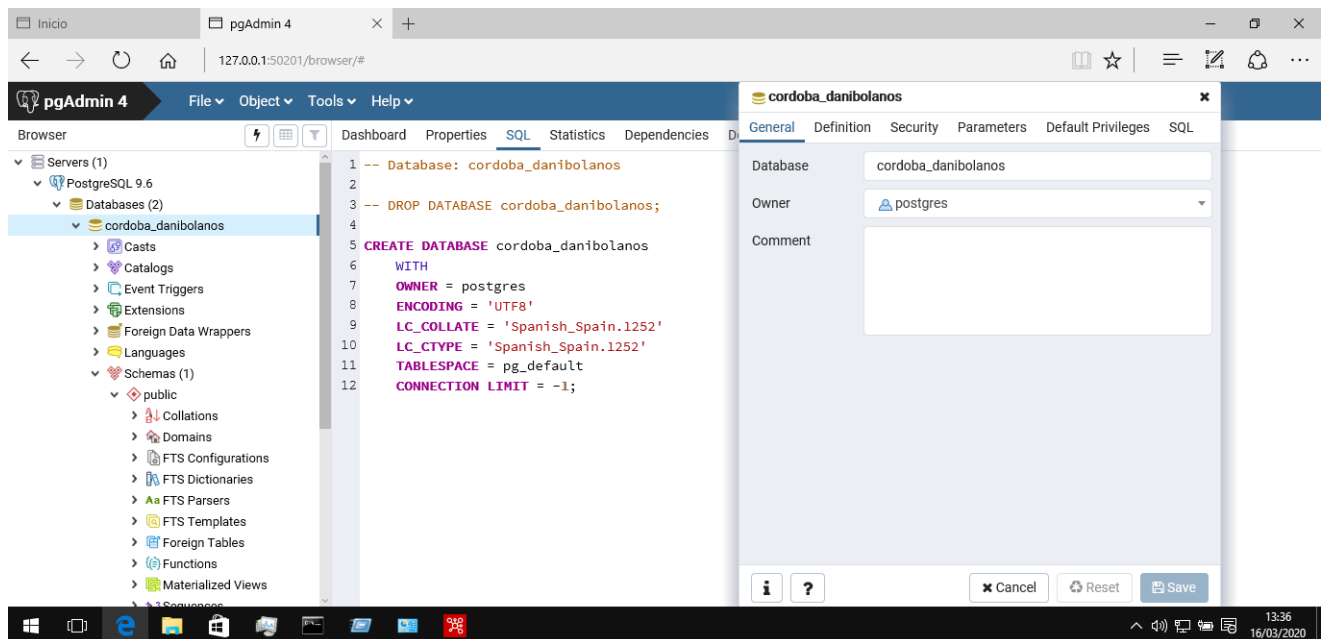


Figura 1: BD creada anteriormente.

Partiremos de la BD **cordoba_danibolanos** creada en la Práctica 3 que contiene las tablas de dimensión **Cuándo** y **Dónde** y la tabla de hechos **Padrón** sobre los datos de la provincia de **Córdoba**.

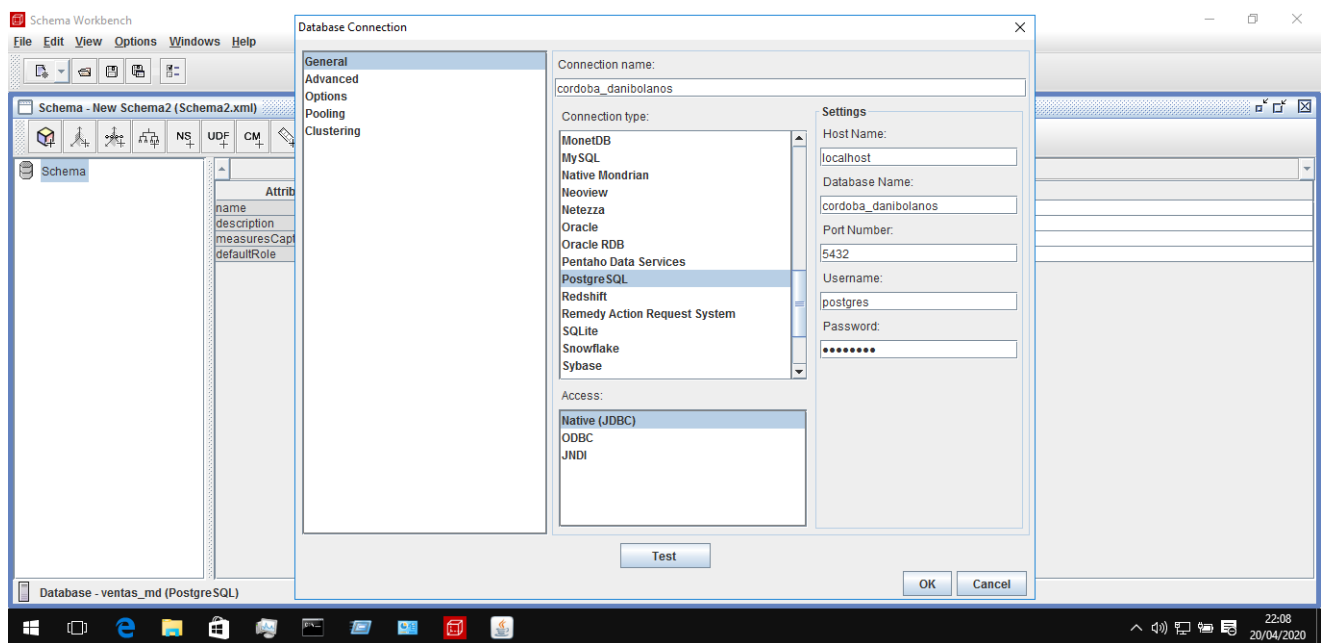


Figura 2: Conexión con la BD.

Pulsamos sobre [«Options», «Connection»] para definir la conexión con la BD que habíamos definido.

En la ventana emergente, seleccionamos el tipo de BD y definimos los parámetros de la conexión como vemos en la imagen.

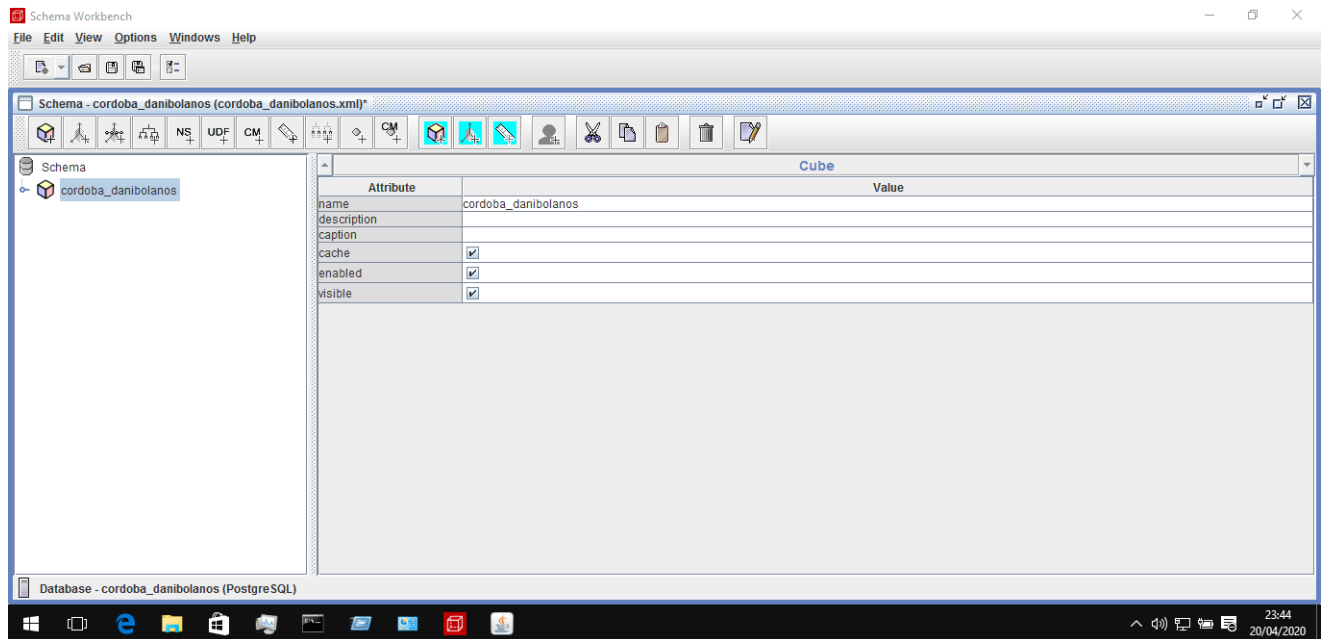


Figura 3: Definición del esquema y el cubo.

Creamos un nuevo esquema con [«File», «New», «Schema»] y definimos su campo 'name' como **cordoba_danibolanos**.

Una vez definido el esquema, dentro de este definimos el cubo tal y como se muestra en la imagen.

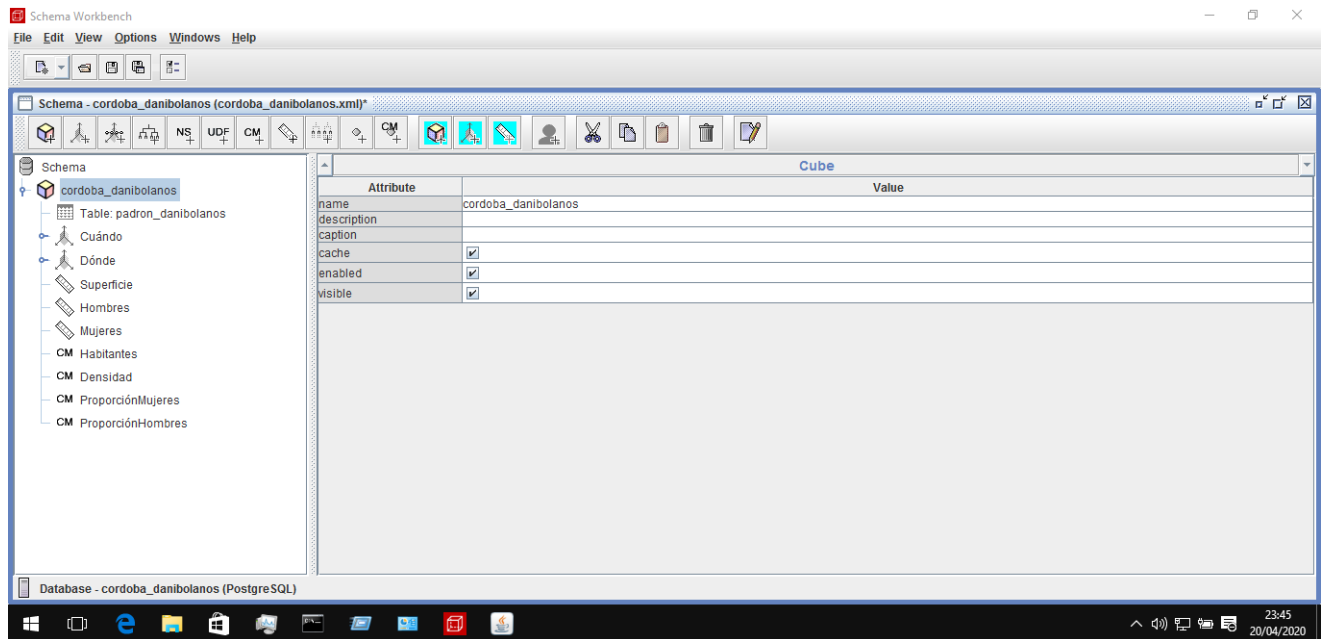


Figura 4: Definición del cubo completa.

He realizado la definición completa del cubo utilizando las diferentes Herramientas proporcionadas por *Schema Workbench*, a continuación se irán mostrando capturas de la definición de cada paso.

En el menú de la izquierda se muestra el resultado final.

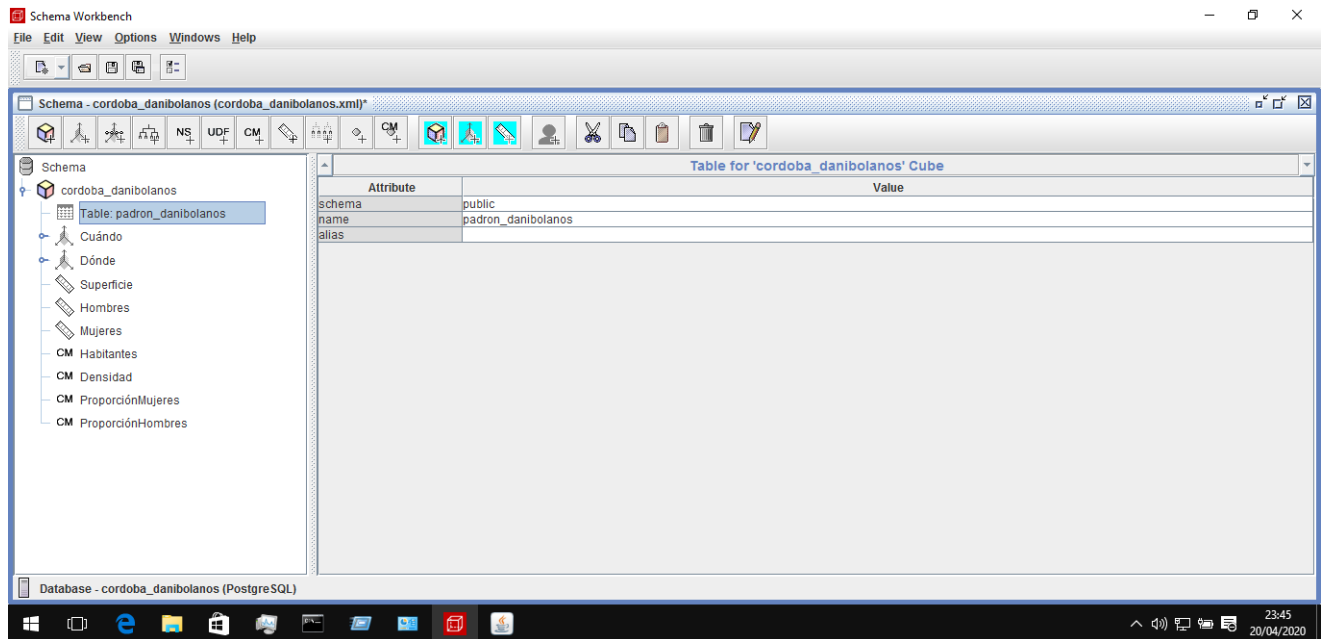


Figura 5: Definición de la tabla de hechos.

Seleccionamos la opción «Add table» para añadir al cubo la tabla de hechos, en nuestro ejemplo se corresponde con la tabla **padron_danibolanos** contenida en el esquema **public** en la BD.

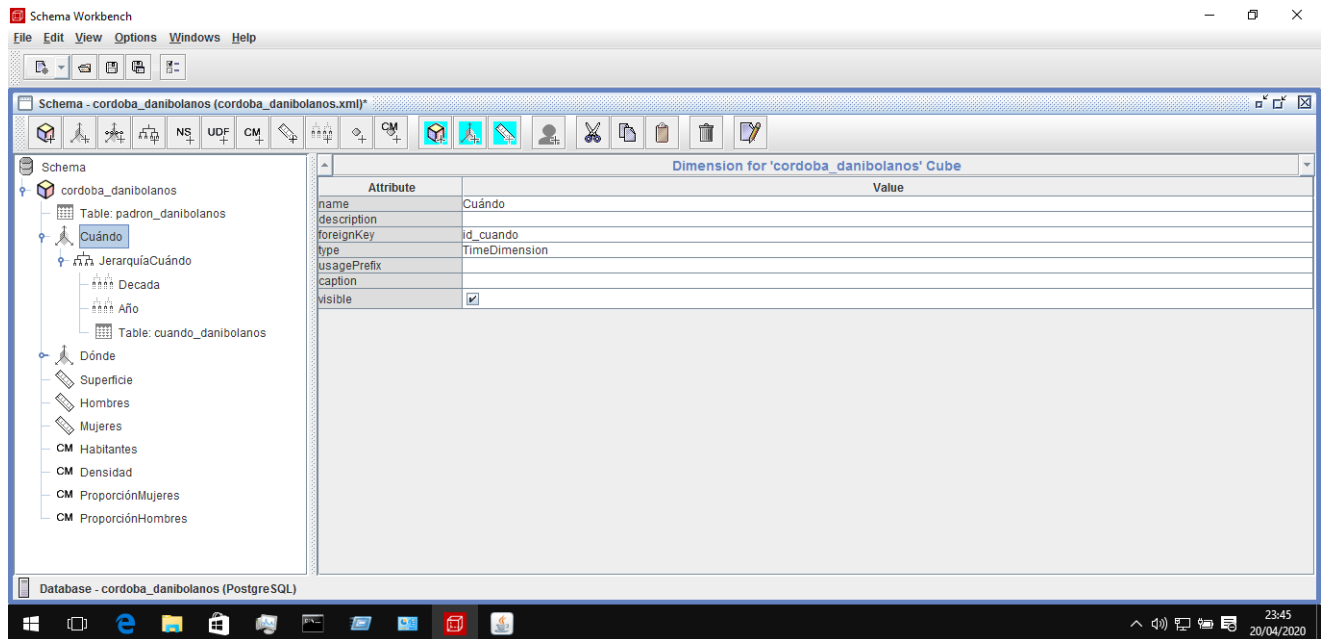


Figura 6: Definición de dimensión **Cuándo**.

Estando situados en el cubo, definimos una nueva dimensión mediante la operación «Add dimension».

Como en este caso la dimensión **Cuándo** es la relativa al tiempo, especificamos su tipo como «TimeDimension».

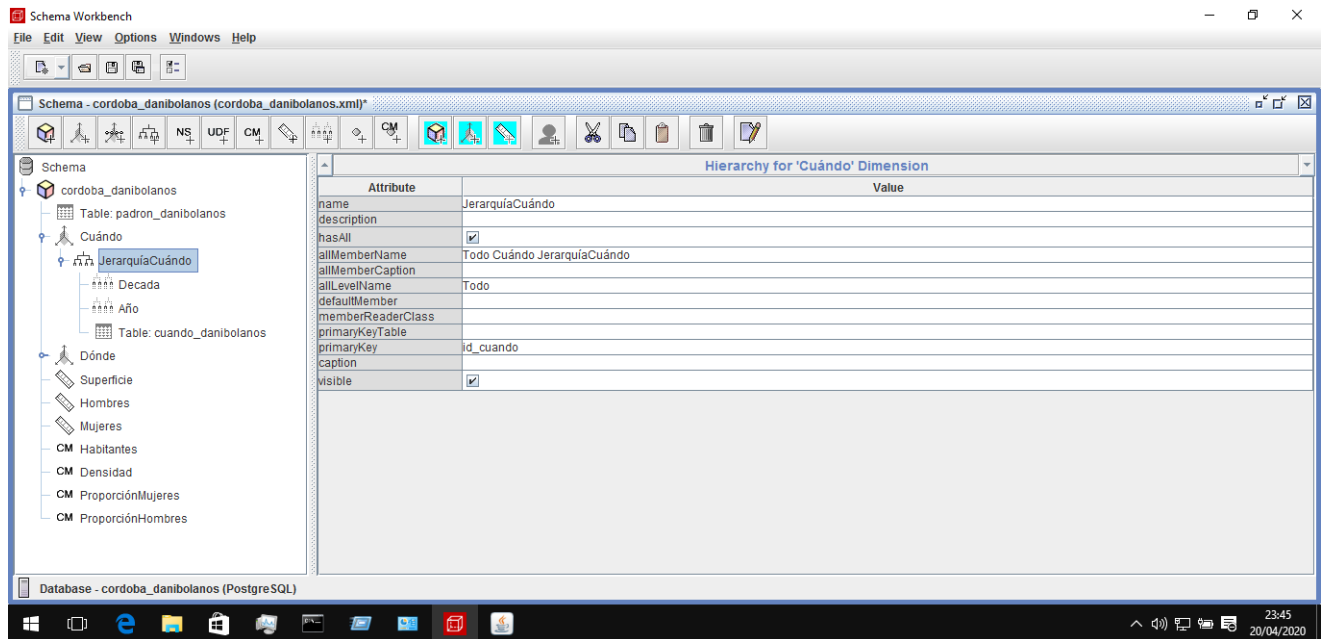


Figura 7: Definición de las jerarquías en **Cuándo**.

A continuación, para cada dimensión definiremos sus jerarquías mediante la operación «Add Hierarchy» desde la dimensión correspondiente.

Nos basaremos en las jerarquías definidas en la práctica 2.2, que en el caso de la dimensión **Cuándo** eran:

- **JerarquíaCuándo: Decada, Año.**

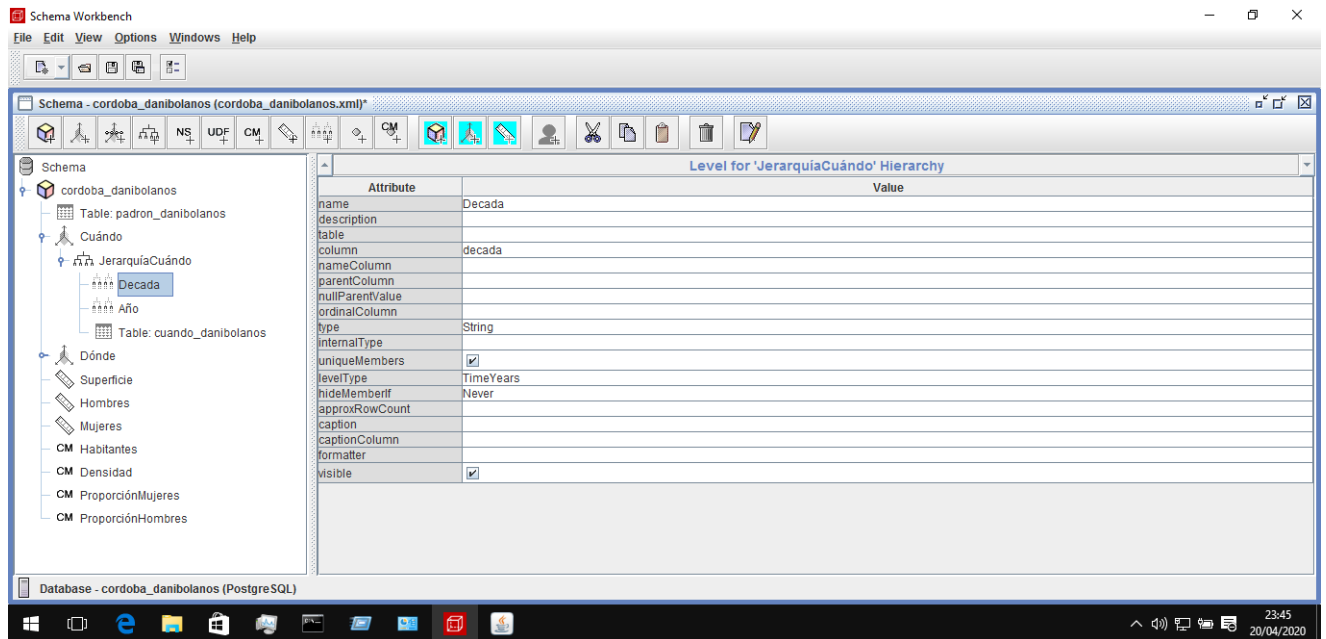


Figura 8: Definición de los niveles de **JearquíaCuándo**.

Cada jerarquía puede tener varios niveles. Se definen mediante la operación «Add Level» desde la jerarquía correspondiente.

En este caso vemos como se define el nivel **Década** donde especificamos sus valores teniendo en cuenta el ejemplo de ventas. Como está contenido en la dimensión relativa al tiempo su **levelType** no puede ser *Regular*, por lo que establecemos el que más se le parezca en este caso *TimeYears*.

El nivel **Año** se define de forma similar, cambiando su nombre y columna por los relativos a año.

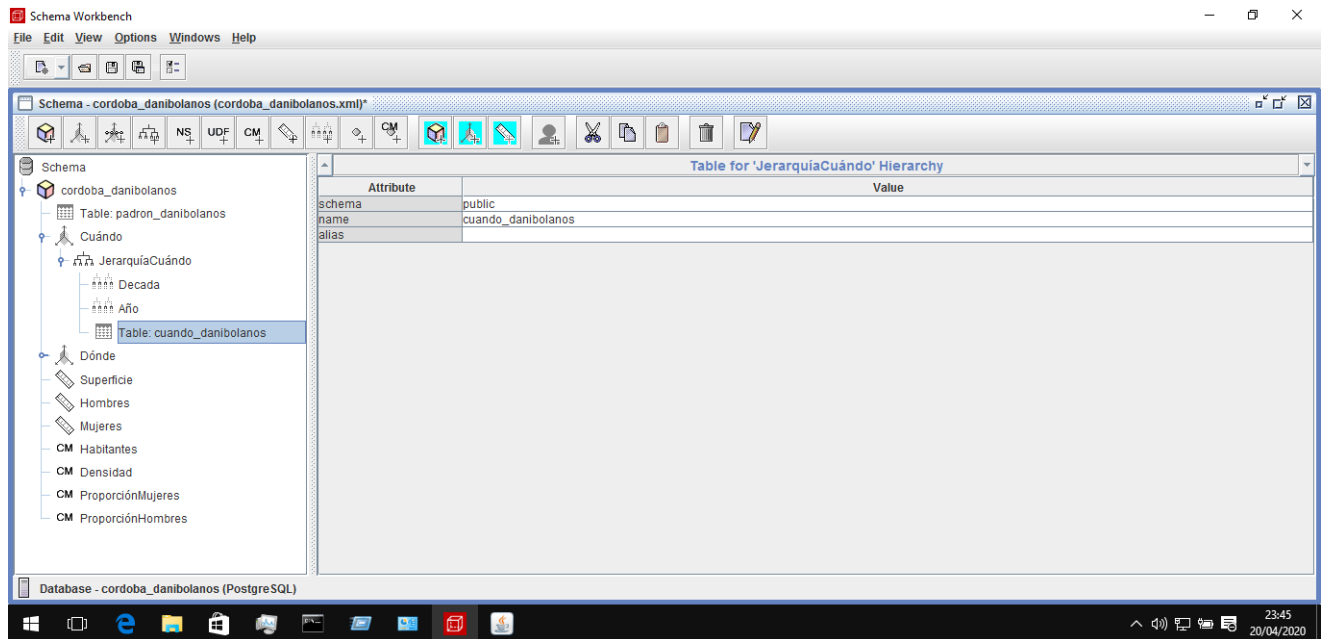


Figura 9: Definición de la tabla **Cuándo** para la jerarquía.

Dentro de cada jerarquía, debemos definir una tabla de referencia a la dimensión correspondiente, en este caso a **Cuándo** de la BD.

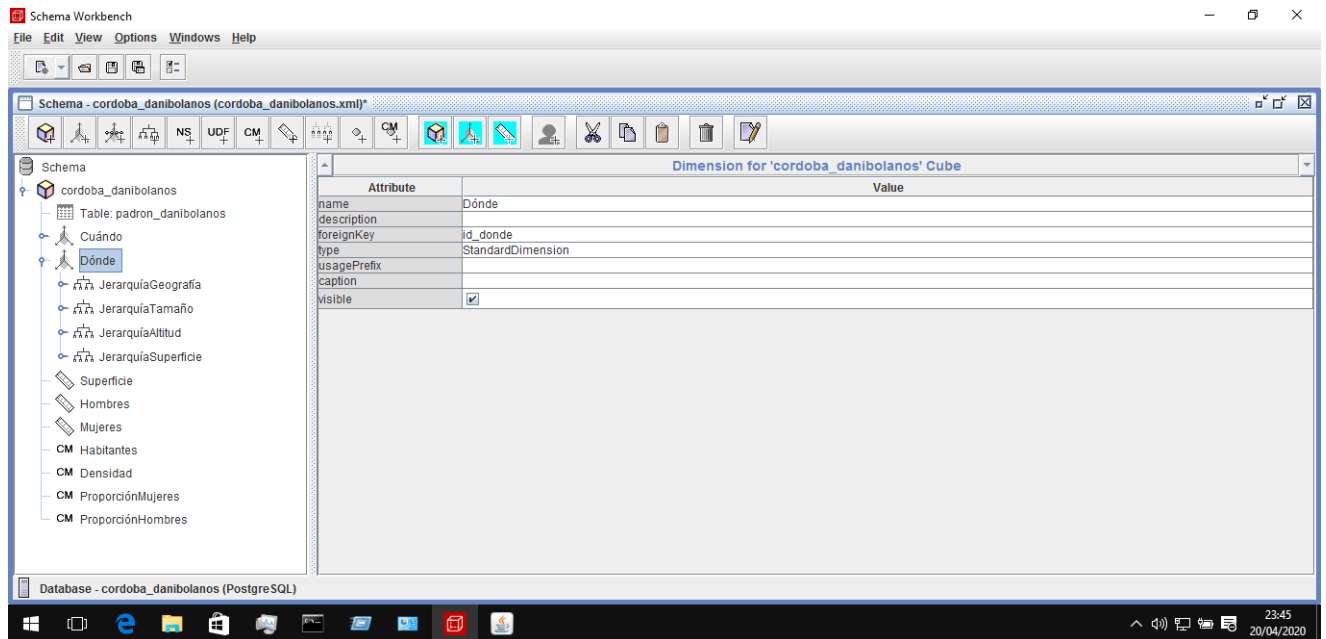


Figura 10: Definición de dimensión **Dónde**.

Estando situados en el cubo de nuevo, definimos la nueva dimensión **Dónde**.

Como en este caso la dimensión **Dónde** no es relativa al tiempo, su tipo será «StandardDimension».

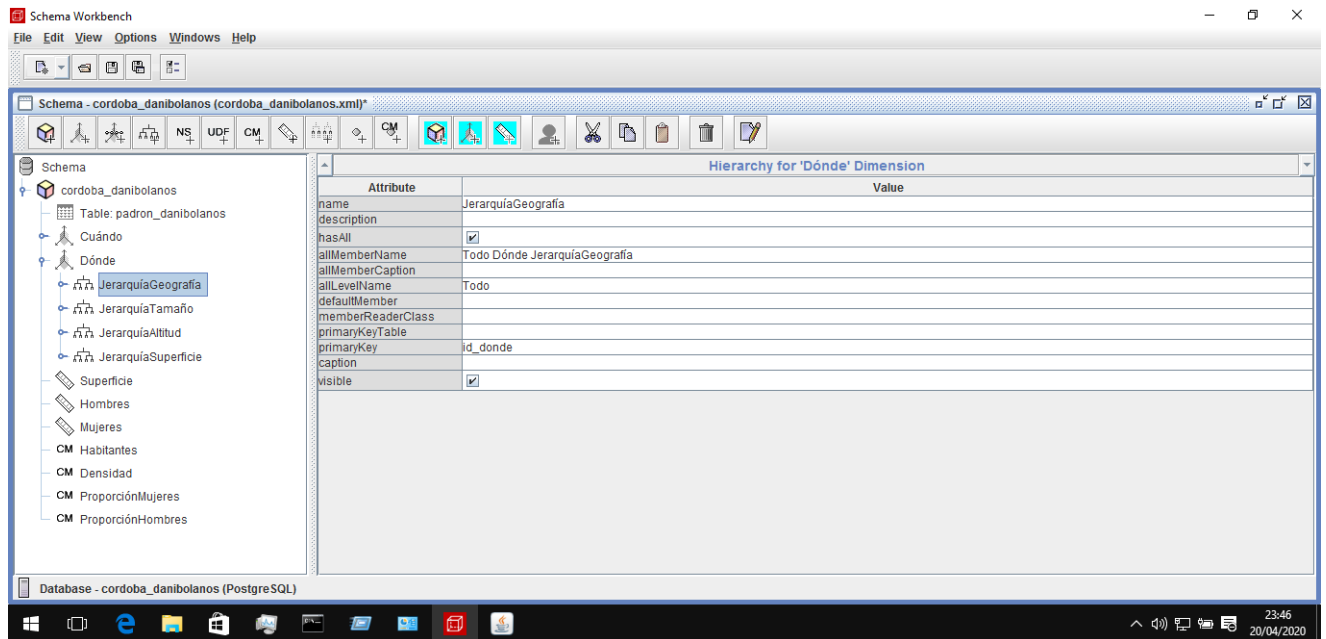


Figura 11: Definición de las jerarquías en **Dónde**.

Nos basaremos en las jerarquías definidas en la práctica 2.2, que en el caso de la dimensión **Dónde** eran:

- **JerarquíaGeografía:** Comarca, Municipio.
- **JerarquíaTamaño:** Tamaño_habitantes, Municipio.
- **JerarquíaAltitud:** Altitud_m, Municipio.
- **JerarquíaSuperficie:** Superficie_Km2, Municipio.

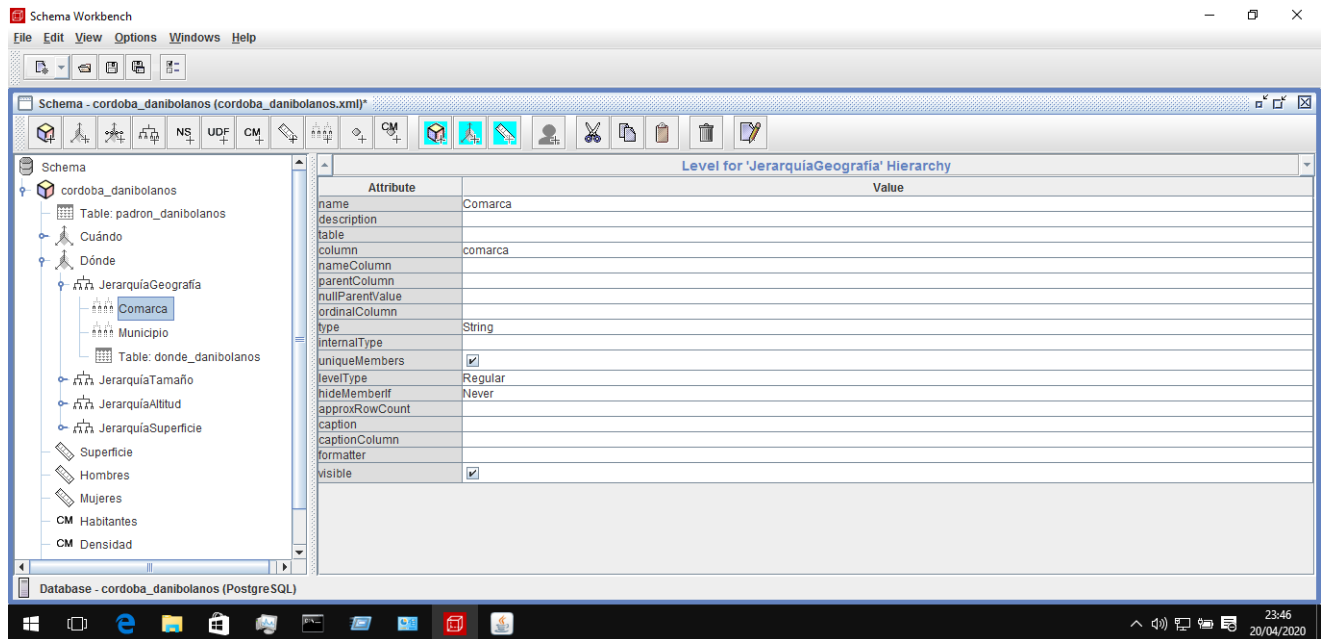


Figura 12: Definición de los niveles de **JerarquíaGeografía**.

En este caso vemos como se define el nivel **Comarca** donde especificamos sus valores teniendo en cuenta el ejemplo anterior, a excepción de que su **levelType** será en este caso *Regular*.

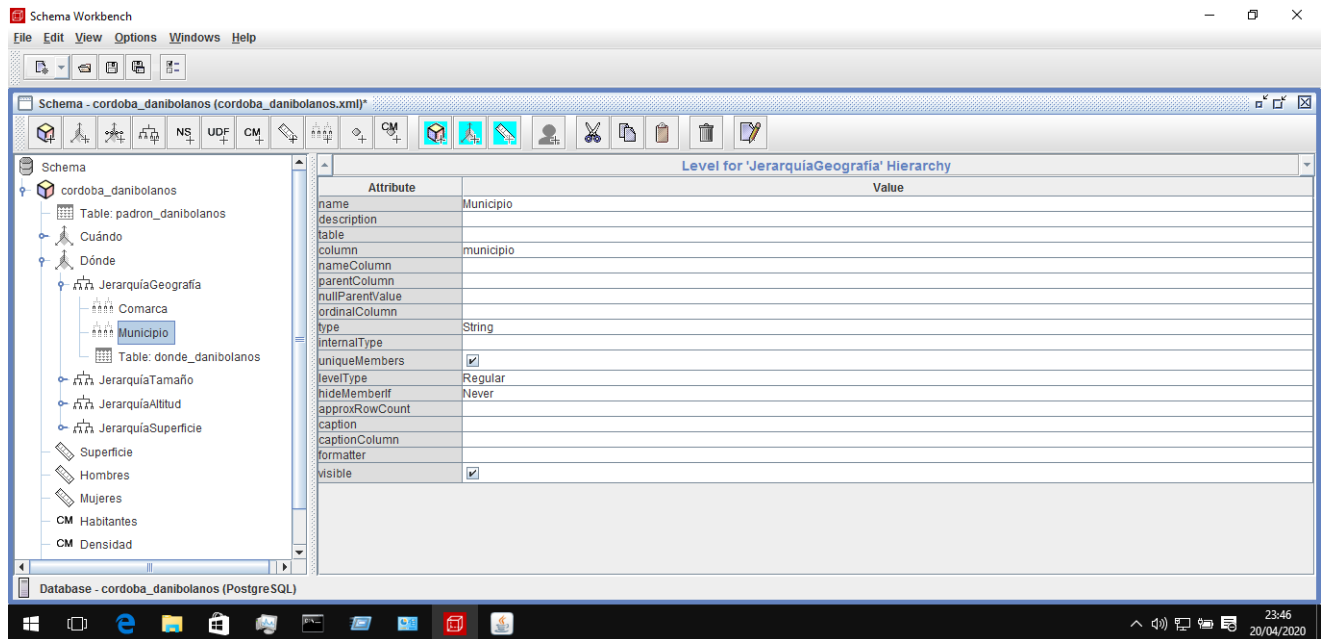


Figura 13: Definición de los niveles de **JerarquíaGeografía**.

Para **Municipio** se hace de forma similar a **Comarca**, especificamos su **levelType** también como *Regular*.

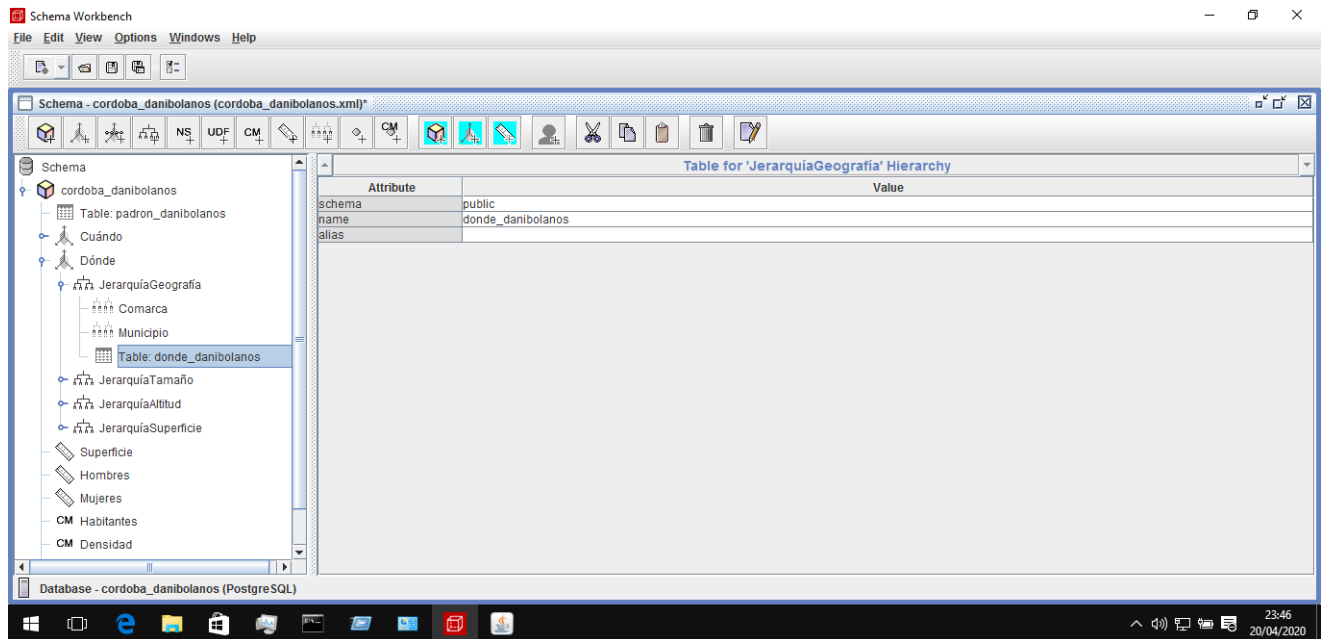


Figura 14: Definición de la tabla **Dónde** para la jerarquía.

Dentro de cada jerarquía, debemos definir una tabla de referencia a la dimensión correspondiente, en este caso a **Dónde** de la BD.

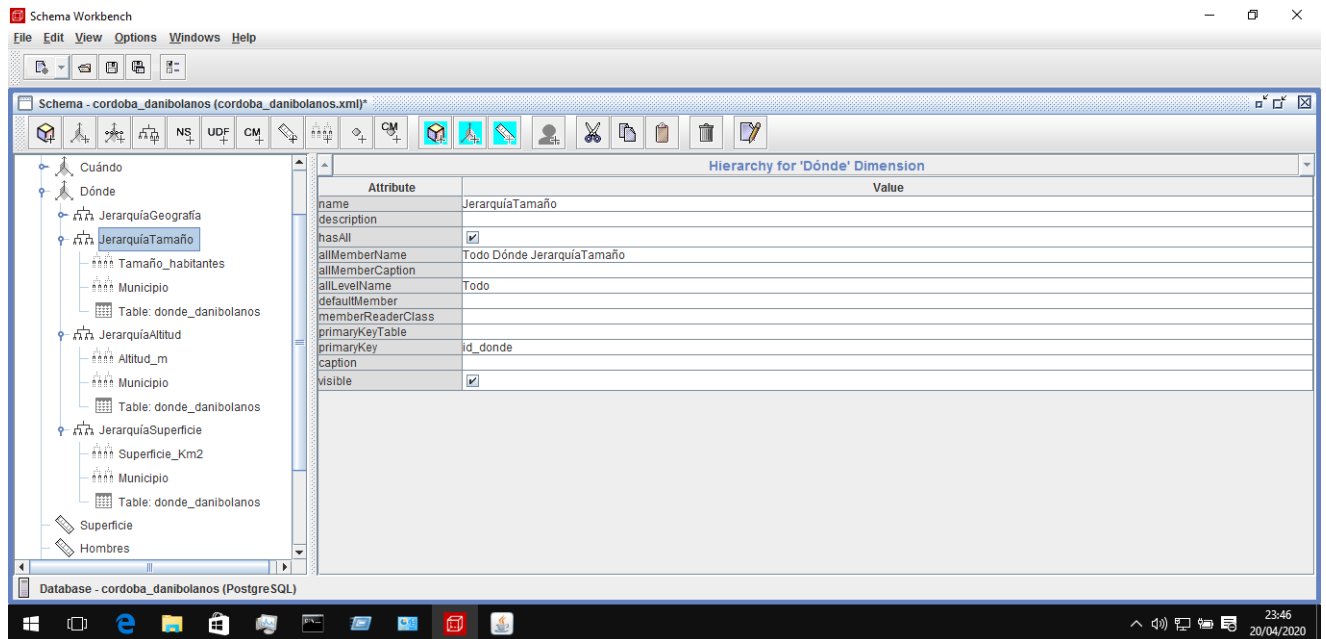


Figura 15: Resto de jerarquías de la dimensión **Dónde**

Para el resto de jerarquías, la definición es similar salvo las referencias a los campos de la tabla (los cuales vienen especificados por el nombre del nivel).

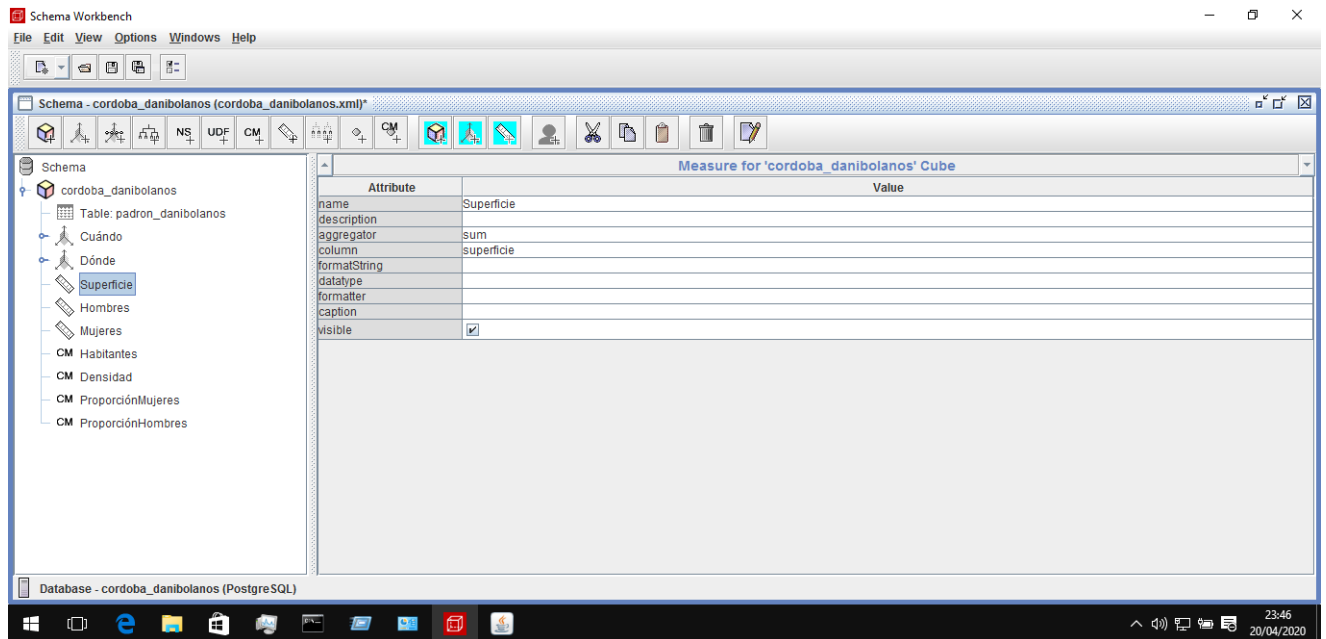


Figura 16: Definición de las mediciones.

Podemos definir mediciones mediante las operaciones «Add Measure» para el cubo.

Definiremos las mediciones **Superficie**, **Hombres** y **Mujeres**, ya que las mediciones calculadas podemos extraerlas desde aquí.

Mostraremos la definición para **Superficie** ya que para el resto de mediciones es similar.

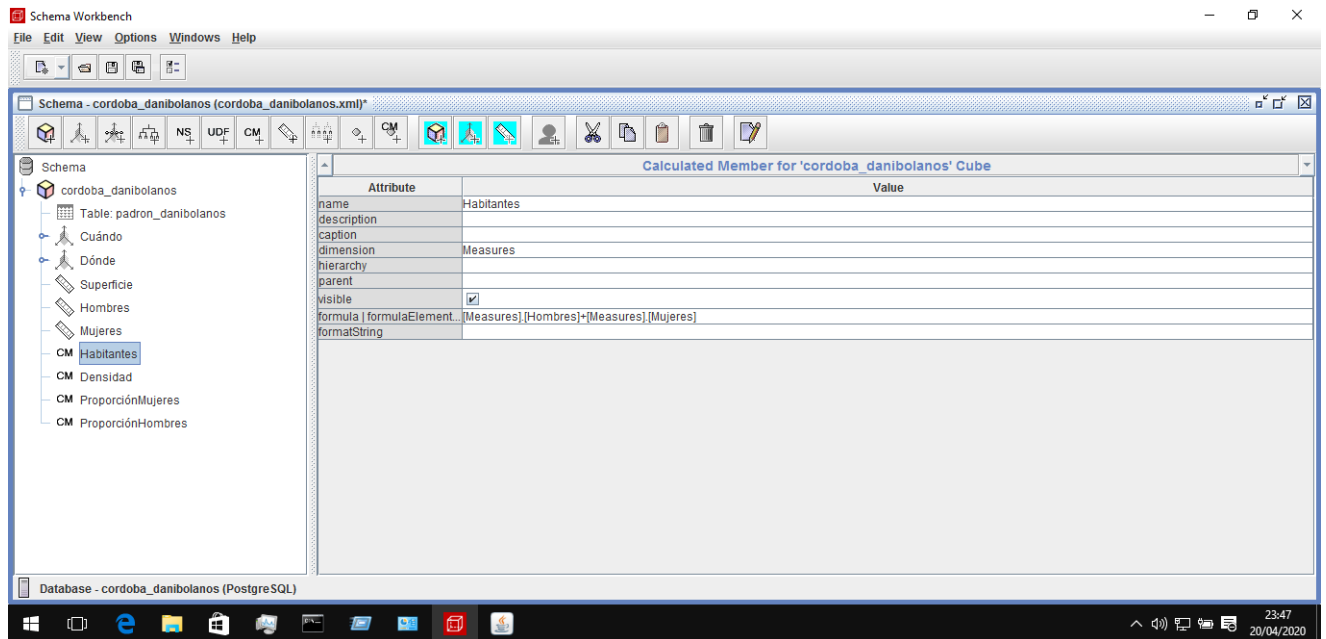


Figura 17: Definición de las mediciones calculadas.

Podemos definir mediciones calculadas mediante las operaciones «Add Calculated Member» para el cubo.

Definiremos las mediciones **Habitantes**, **Densidad** y **Promedio Hombres/Mujeres** (solo mostraremos la definición de uno por la dualidad de la fórmula).

Para definir las fórmulas lo hacemos adaptando la fórmula de la práctica 2.2 al lenguaje XML.

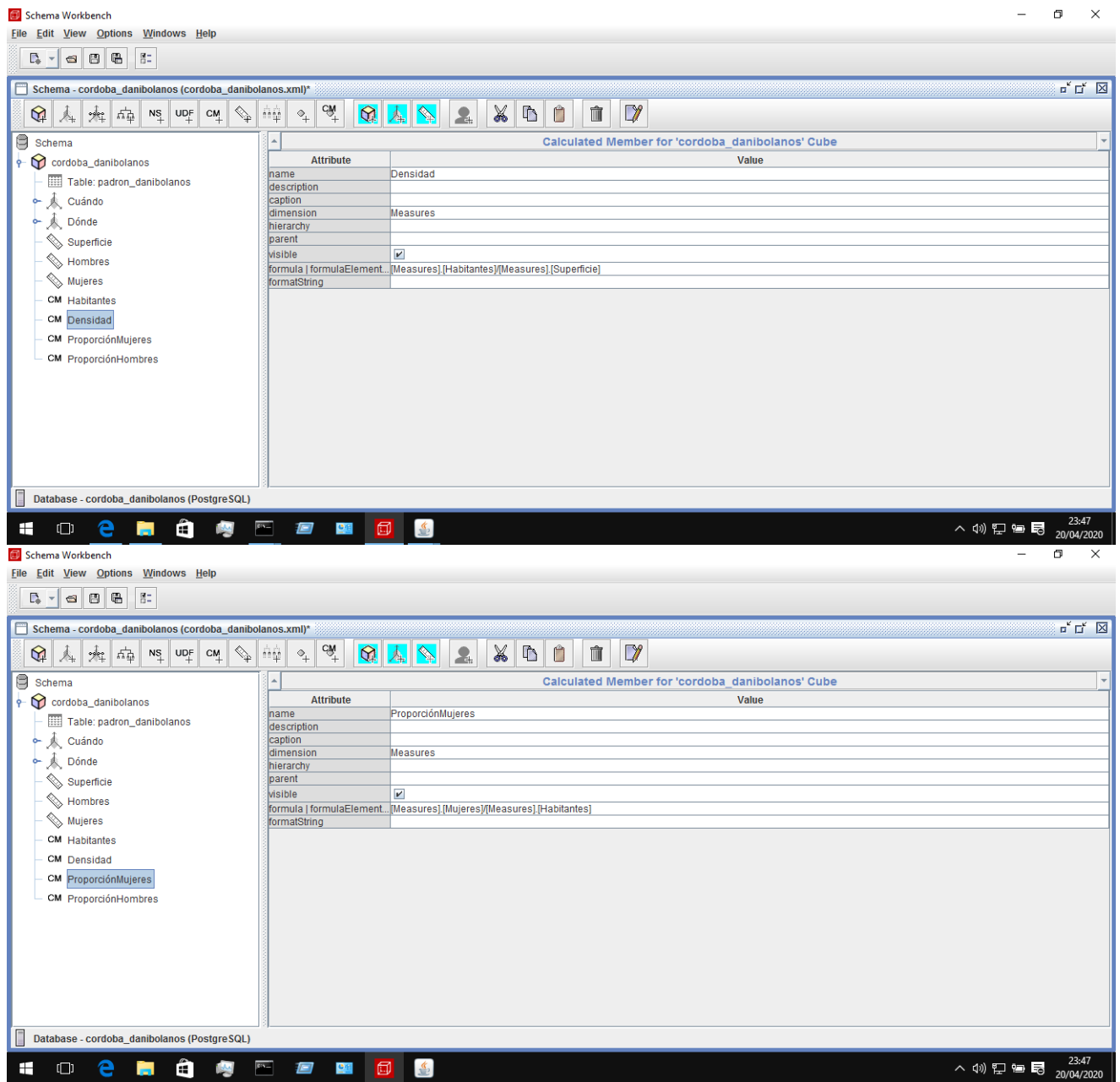


Figura 18: Definición de las mediciones calculadas.

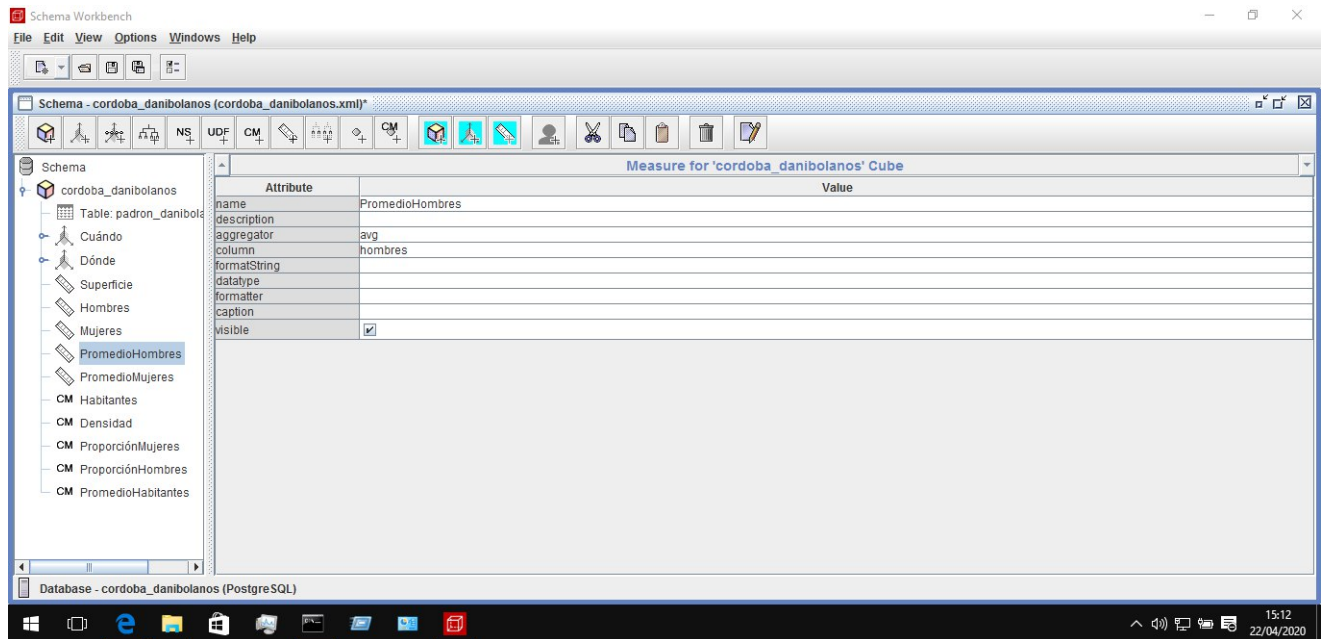


Figura 19: Definición medida **PromedioHombres**.

También podemos definir los **Promedios** para **Hombres** y **Mujeres** indicando en el campo aggregator la opción *avg* (average).

Definiremos **Promedio Hombres** y **Promedio Mujeres** y finalmente calcularemos la medida **Promedio Habitantes** como suma de ambas medidas.

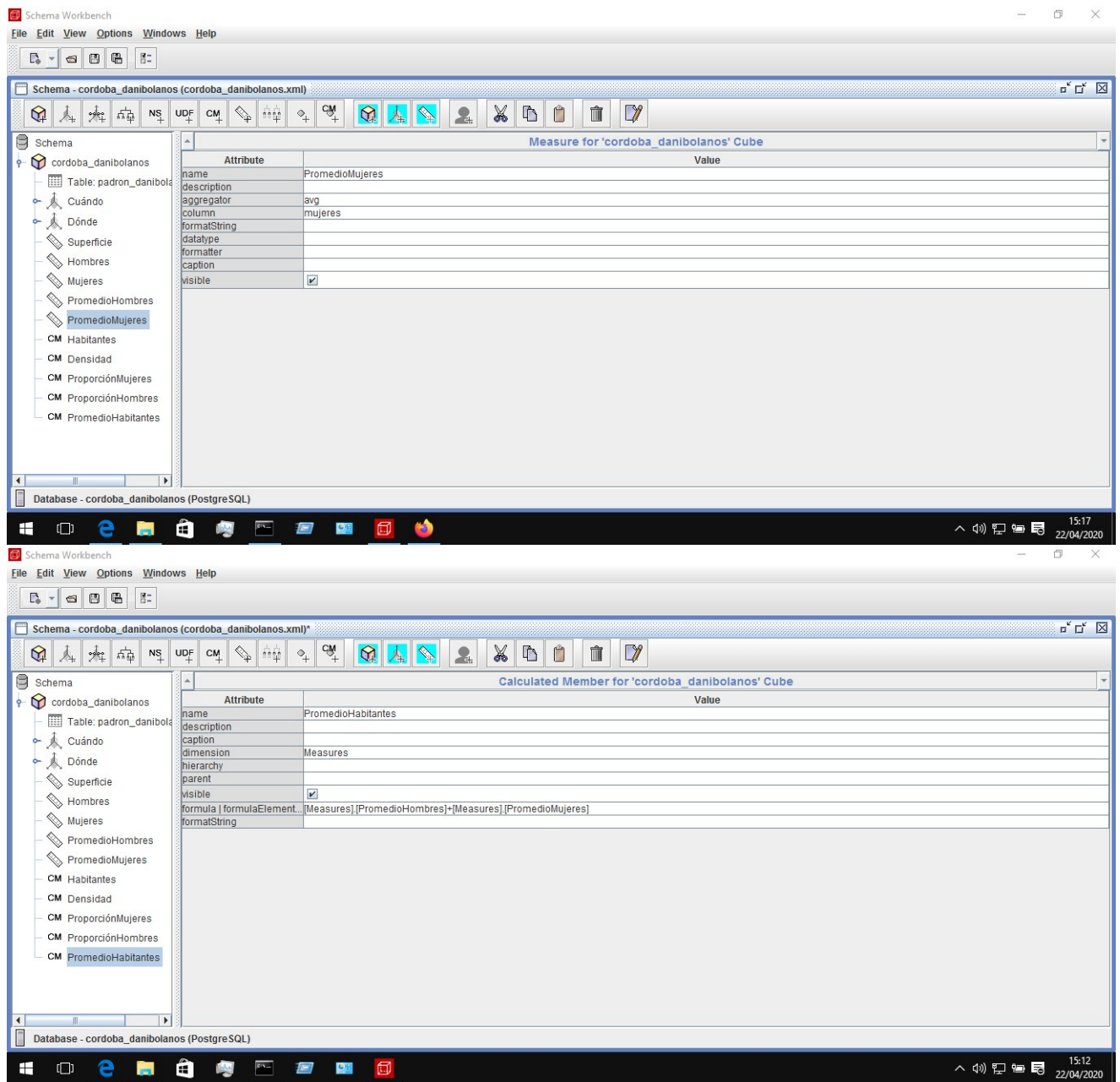


Figura 20: Definición medidas promedios.

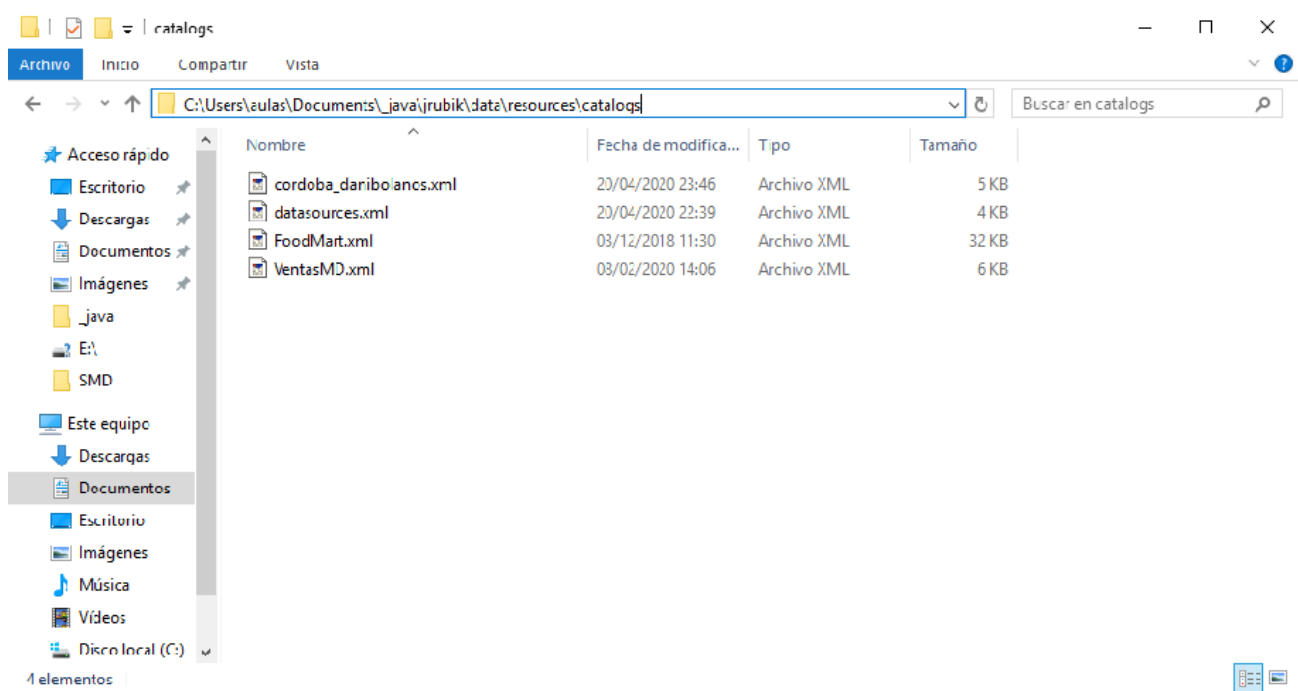


Figura 21: Guardamos el archivo XML generado.

Generamos el archivo XML y lo guardamos en la carpeta de catálogos en la ruta relativa `_java/jrubik/data/resources/catalogs` con el nombre de **cordoba_danibolanos.xml**.

2. Sobre Mondrian, utilizando el esquema multidimensional:

Procedemos a abrir el programa *JRubik* para abrir el archivo y realizar las consultas OLAP en *Mondrian*.

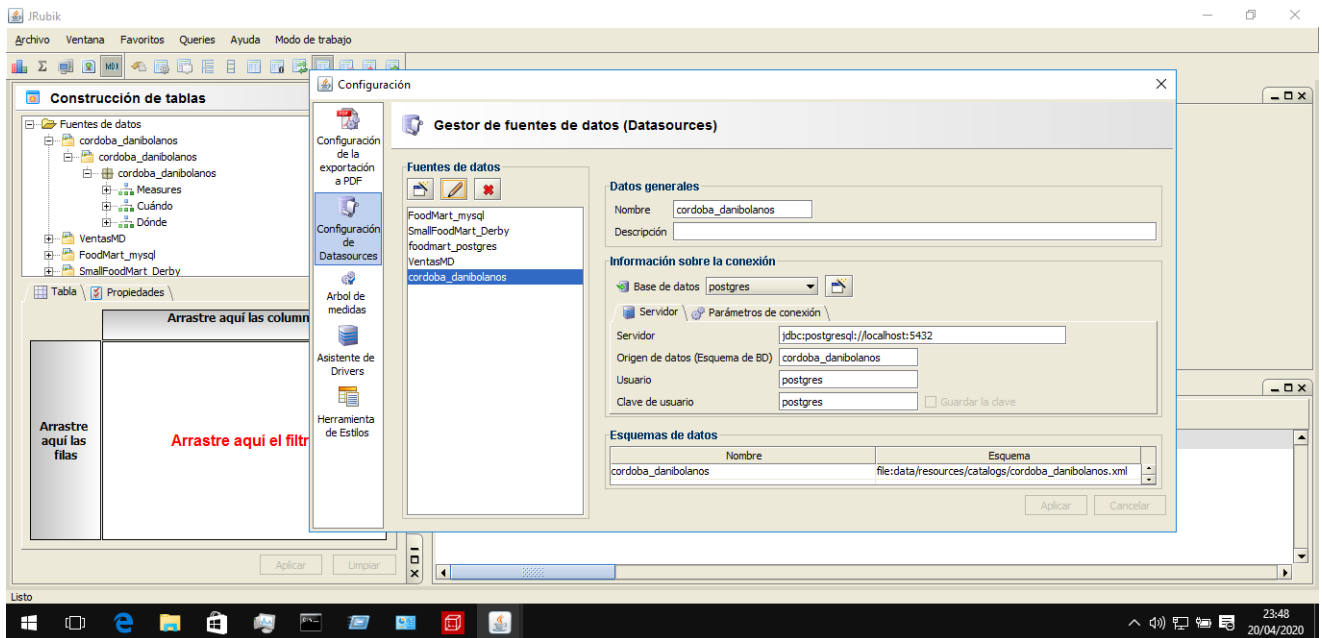


Figura 22: Definición de la fuente de datos.

Definimos la fuente de datos a partir de la BD de *PostgreSQL* y el archivo XML que hemos creado en el ejercicio anterior.

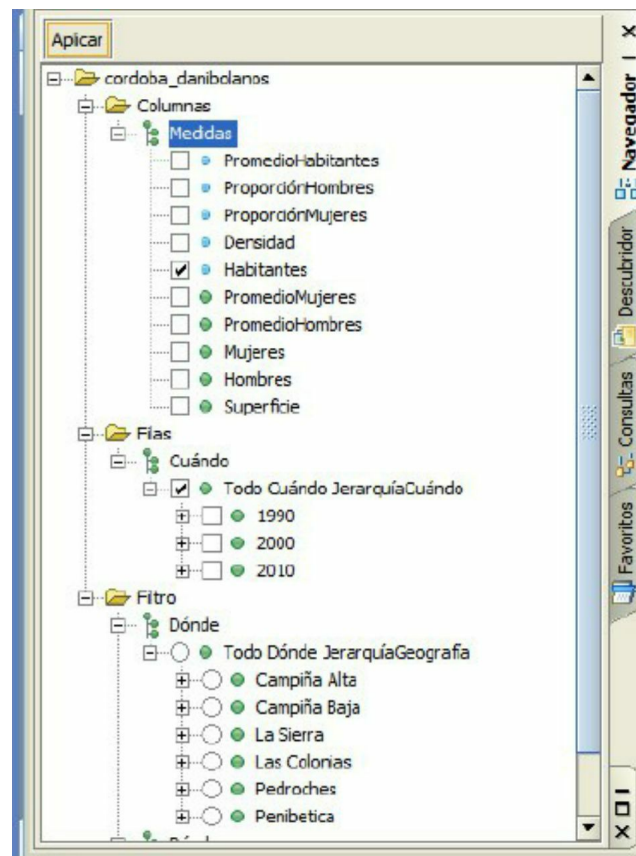
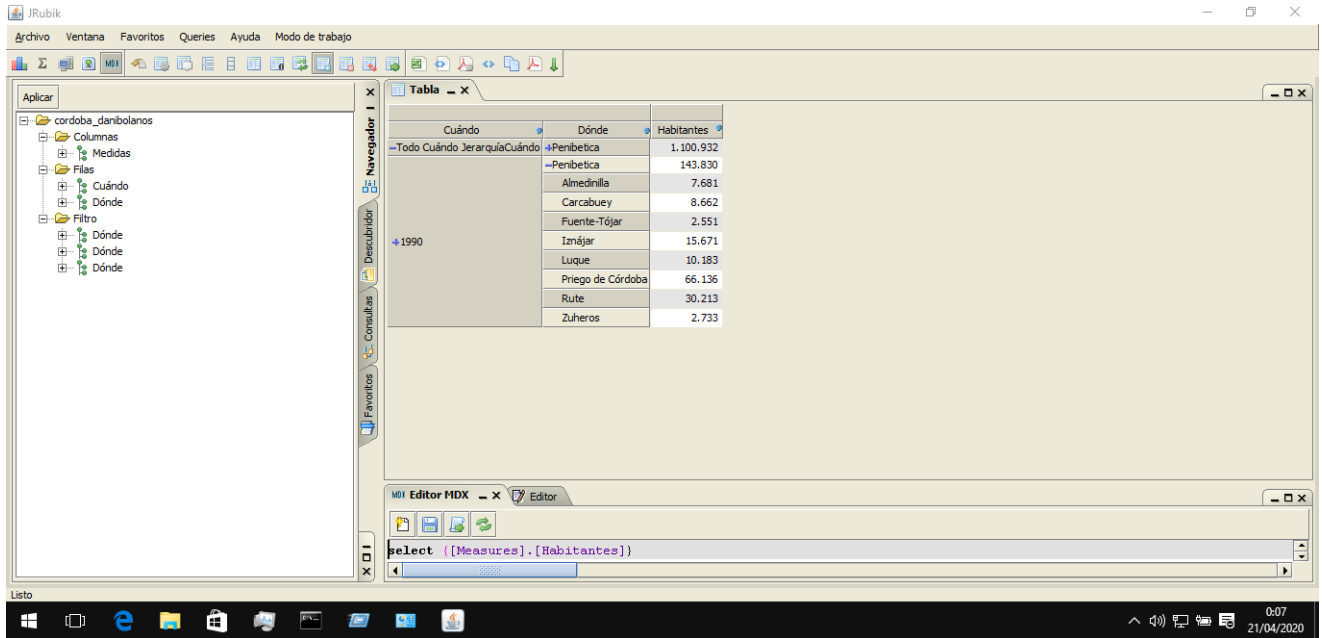


Figura 23: Ventana Navegador en JRubik.

En la ventana 'Navegador' podemos refinar la construcción de la tabla. Permite seleccionar elementos adicionales o bien elementos distintos de los componentes incluidos en las filas, columnas y filtro.

Podemos ver todas las jerarquías y medidas definidas en el archivo XML.

2.1. Obtén un informe inicial libre y explica su contenido.



Cuándo	Dónde	Habitantes
Todo Cuándo JerarquíaCuándo	+Penibética	1.100.932
	+Penibética	143.830
	+Almedinilla	7.681
	+Carcabuey	8.662
	+Fuente-Tójar	2.551
	+Iznájar	15.671
	+Luque	10.183
	+Priego de Córdoba	66.136
	+Rute	30.213
	+Zuñeros	2.733

Figura 24: Informe inicial libre.

El valor de la medición que estamos utilizando es el de la medición calculada **Habitantes**.

Contenido del informe: Número de Habitantes de los municipios de la comarca Penibética de la provincia de Córdoba en la década de 1990.

Para cada dimensión nos encontramos en los siguientes niveles:

- Nivel **Década** en la dimensión **Cuándo** → Restringido a la década de 1990.
- Nivel **Municipio** en la dimensión **Dónde** → Restringido a municipios de la comarca Penibética.

2.2. Mediante Drill-down obtén un nuevo informe y explica su contenido.

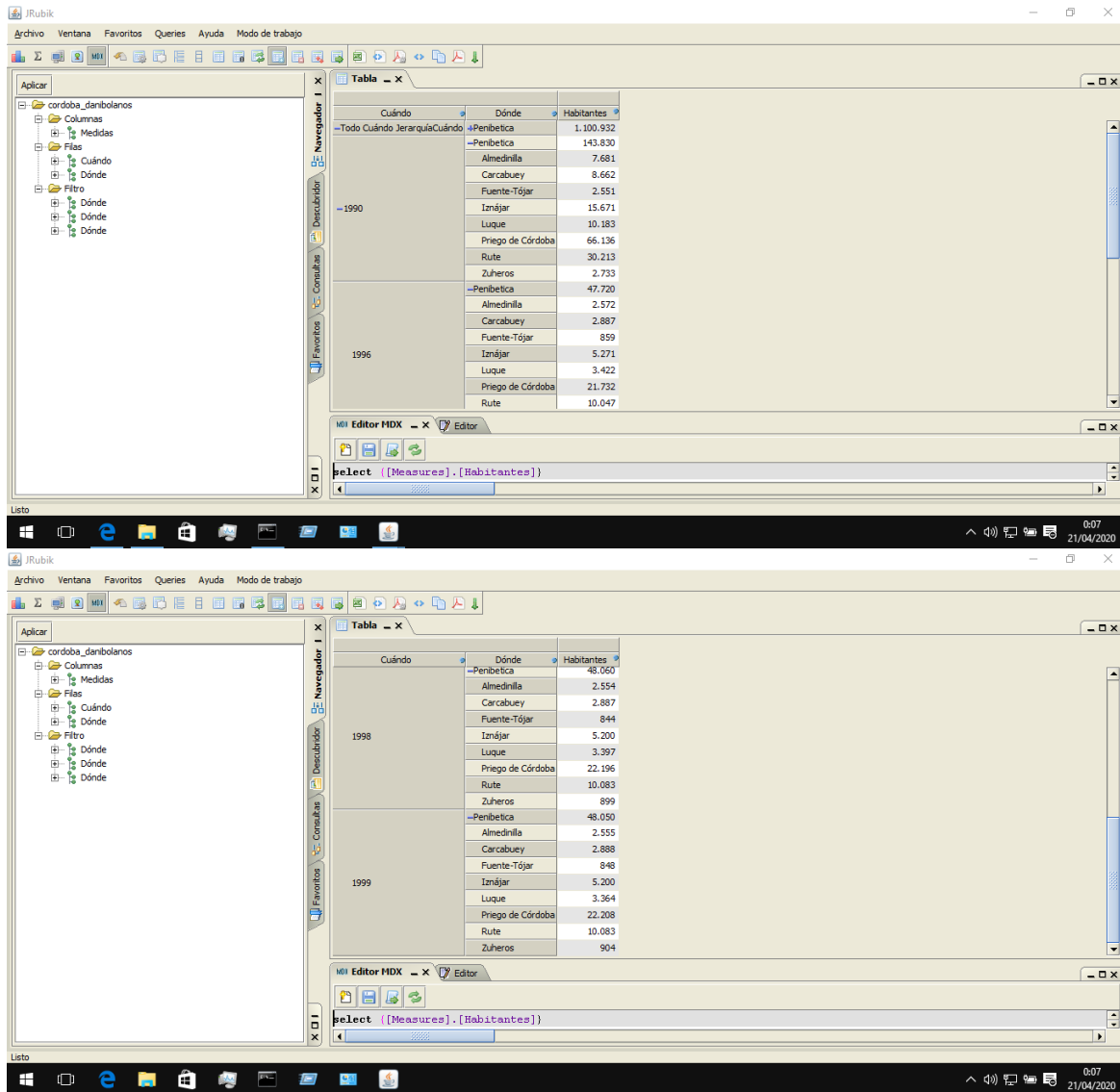


Figura 25: Operación de Drill-down.

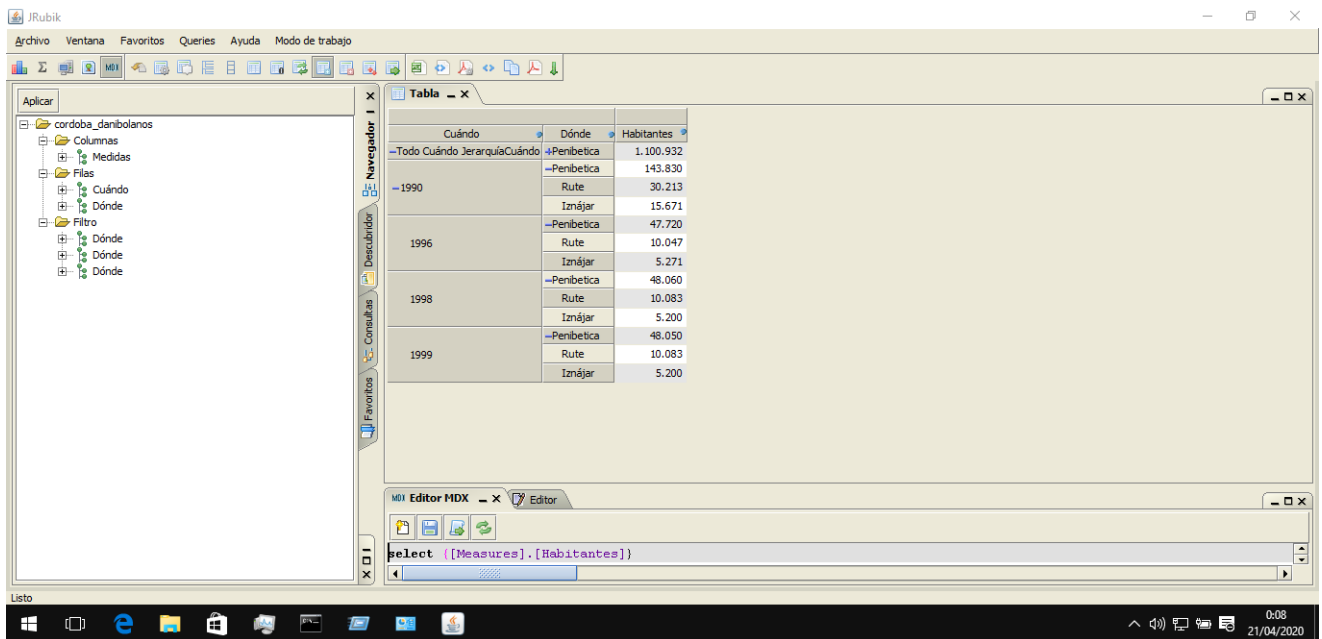
Realizamos una operación de **Drill-down** por la dimensión **Cuándo** para pasar desde el nivel **Década** a **Año**.

Contenido del informe: Número de Habitantes de los municipios de la comarca Penibética en los años 1996, 1998 y 1999.

Para cada dimensión nos encontramos en los siguientes niveles:

- Nivel **Año** en la dimensión **Cuándo** → Restringido a los años de la década de 1990 (1996, 1998 y 1999).
- Nivel **Municipio** en la dimensión **Dónde** → Restringido a municipios de la comarca Penibética.

2.3. Mediante Slice&Dice obtén un nuevo informe y explica su contenido.



Cuándo	Dónde	Habitantes
→ Todo Cuándo Jerarquía Cuándo	→ Penibetica	1.100.932
	→ Rute	143.830
→ 1990	→ Rute	30.213
	→ Iznájar	15.671
	→ Penibetica	47.720
→ 1996	→ Rute	10.047
	→ Iznájar	5.271
	→ Penibetica	48.060
→ 1998	→ Rute	10.083
	→ Iznájar	5.200
	→ Penibetica	48.050
→ 1999	→ Rute	10.083
	→ Iznájar	5.200

MDX Editor MDX Editor

```
select ([Measures].[Habitantes])
```

Figura 26: Operación de **Slice&Dice**.

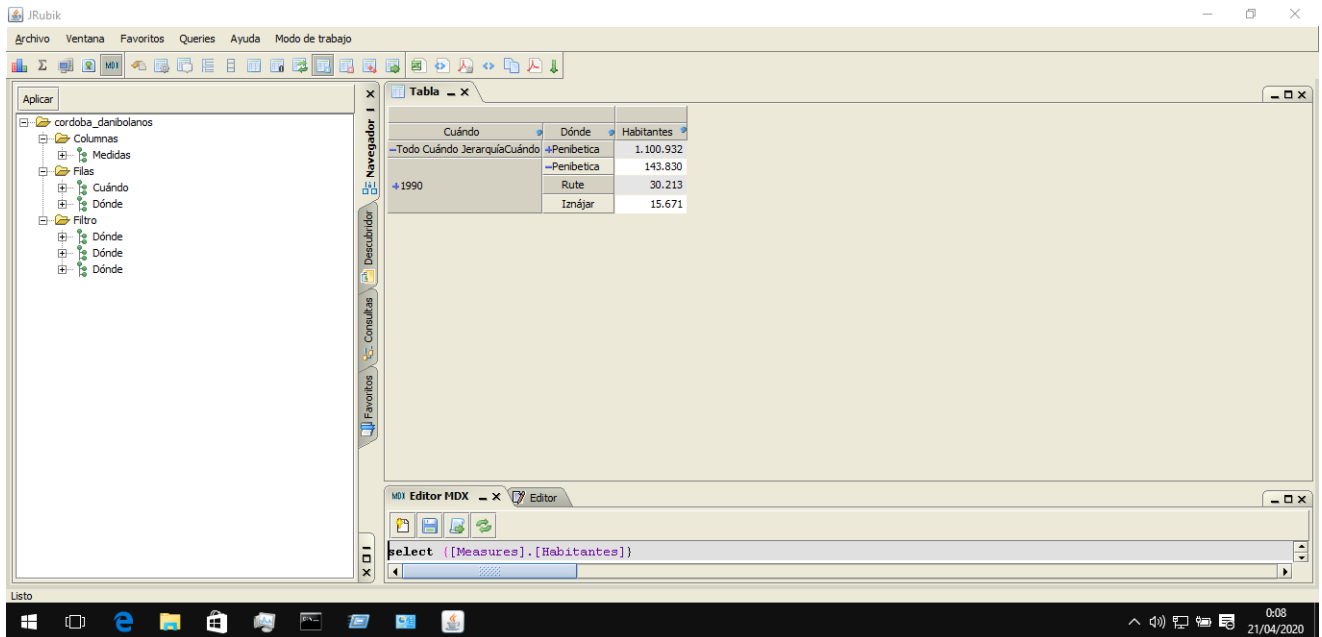
Realizamos una operación de **Slice&Dice** para restringir el campo **Municipio** a los municipios de **Rute** e **Iznájar**.

Contenido del informe: Número de Habitantes de los municipios de Rute e Iznájar en los años 1996, 1998 y 1999.

Para cada dimensión nos encontramos en los siguientes niveles:

- Nivel **Año** en la dimensión **Cuándo** → Restringido a los años de la década de 1990 (1996, 1998 y 1999).
- Nivel **Municipio** en la dimensión **Dónde** → Restringido a municipios de Rute e Iznájar.

2.4. Mediante Roll-up obtén un nuevo informe y explica su contenido.



Cuándo	Dónde	Habitantes
→ Todo Cuándo Jerarquía Cuándo	→ Penibetica	1.100.932
→ 1990	→ Rute	143.830
	→ Iznájar	30.213
		15.671

MDX Editor MDX Editor

```
select ([Measures].[Habitantes])
```

Figura 27: Operación de **Roll-up**.

Realizamos una operación de **Roll-up** por la dimensión **Cuándo** para pasar desde el nivel **Año** a **Década**.

Contenido del informe: Número de Habitantes de los municipios de Rute e Iznájar en la década de 1990.

Para cada dimensión nos encontramos en los siguientes niveles:

- Nivel **Década** en la dimensión **Cuándo** → Restringido a la década de 1990.
- Nivel **Municipio** en la dimensión **Dónde** → Restringido a municipios de Rute e Iznájar.

3. Bibliografía.

Referencias

- [1] Guión de prácticas de la asignatura.