

# PROPUESTA DE IMPLEMENTACIÓN DE UN WAREHOUSE CON EXTRACCIÓN DE DATOS DE UNA BASE DE DATOS ESPACIAL Y RELACIONAL PARA LA GESTIÓN TERRITORIAL



PROPOSAL FOR IMPLEMENTING A DATA WAREHOUSE WITH EXTRACTION OF A DATABASE TO RELATIONAL SPACE AND TERRITORIAL MANAGEMENT

**AUTOR**

ANGÉLICA URRUTIA SEPÚLVEDA  
DOCTORA  
Universidad Católica del Maule  
aurrutia@ucm.cl  
CHILE

**AUTOR**

REINALDO MEDINA  
Magister ©  
Universidad Santiago de Chile  
rmedina@gmail.com  
CHILE

**INSTITUCIÓN**

Universidad Católica del Maule  
Casa centra, Avenida San Miguel 3605  
Teléfono. (56) (71) 203 100  
info@ucm.cl  
TALCA, CHILE

**INSTITUCIÓN**

USACH  
Avenida Libertador, Bernardo O'Higgins  
3363 Estación Central  
(562)7180000  
SANTIAGO, CHILE

**Recepción:** Junio 30 de 2009

**Aceptación:** Agosto 14 de 2009

**Temática:** Gestión tecnológica

**Tipo de Artículo:** Artículo de investigación científica y tecnológica

**RESUMEN**

Este trabajo presenta un conjunto de pasos para la creación de un Data Warehouse Espacial (DWE), de forma sencilla y amigable, que provea las bases de conocimiento a los diseñadores para que puedan construir implementaciones similares. Se consideran como fuente de datos de extracción bases de datos relacionales y bases de datos espaciales, que generan un esquema global multidimensional asociado a un DWE. Dentro de sus etapas más relevantes esta la definición de especificaciones de requerimiento e indicadores de gestión, construcción de modelos conceptuales de cubos, diseño lógico de estrella o constelación, mapeo de datos para la construcción de cubos desde los datos fuentes, para finalmente implementar y consultar un Data Warehouse Espacial. Todo ello aplicado a un caso práctico de la comuna de Puente Alto, exhibiendo los resultados obtenidos en este estudio. Lo aquí lo expuesto es una forma de construir un sistema DWE de apoyo a la toma de decisiones.

## PALABRAS CLAVES

Gestión de datos Espacial  
 indicadores de gestión Territorial  
 Data Warehouse Espacial

## ABSTRACT

This paper presents a set of steps for creating a Data Warehouse Space (DWE), a simple and friendly, providing knowledge bases for designers so they can build similar implementations. Are considered as a source of data mining relational databases and spatial databases, generating an overall multidimensional schema associated with a similar implementations DWE. Among its most important stages is the definition of requirement specifications and management indicators, construction of conceptual models of cubes, logical design of a star or snowflake, mapping data to build cubes from the data sources, and ultimately implement and consult a Spatial Data Warehouse. This applied to a case study from the municipality of Puente Alto, displaying the results obtained in this study. What is described here as a way to build a DWE system to support decision making.

## KEYWORDS

Spatial data management  
 indicators of Territorial Management  
 Data Warehouse Space

## INTRODUCCIÓN

El Data Warehouse Espacial (DWE) entrega datos geo-descriptivos, un aporte interesante en su implementación es contar con una metodología de construcción que describa paso a paso sus requerimientos. Todo ello, porque en su creación se requiere un acabado análisis de cada uno de los componentes que intervienen directamente en cada proceso, nos referimos a un estudio profundo de las bases de datos espaciales y los conceptos de Data Warehouse Espacial, material que sirven de plataforma para estructurar los procedimientos, los procesos e indicadores que deben ser tomados en cuenta cuando existe interés en este tipo de iniciativa.

Se debe tener en cuenta la complejidad de preparar cada uno de los pasos que permite crear un DWE, dado que la estructuración de la información es uno de los procesos más complejos, además de que la data para generar este tipo de implementación suele requerir bases de datos diferentes y por lo tanto los datos no tienen coincidencias (datos de coordenadas y datos relacionales).

Se publica en (Berkel, 2003) en la página Web de Esri (Empresa de Software Geográficos) en forma breve, una definición de lo que es un DWE y cuales son los resultados que podrían esperarse en un estudio utilizando este instrumento. Proponen que un DWE debe ser planificado de acuerdo al ámbito de negocio en que será aplicado, estudiando la calidad y cantidad de información utilizada. El autor describe las componentes que conforman un Data Warehouse Espacial y las características básicas de los datos. Otros trabajos similares se encuentran en Doglas y Nerbert, 1994), (Petraikos, 2002), (Terri et al, 2002).

En (Peralta, 2001) se describe una metodología de diseño y mapeo desde una base de datos fuente a una base de datos multidimensional y sus operaciones de navegación las que son utilizadas en este trabajo en la creación del DWE.

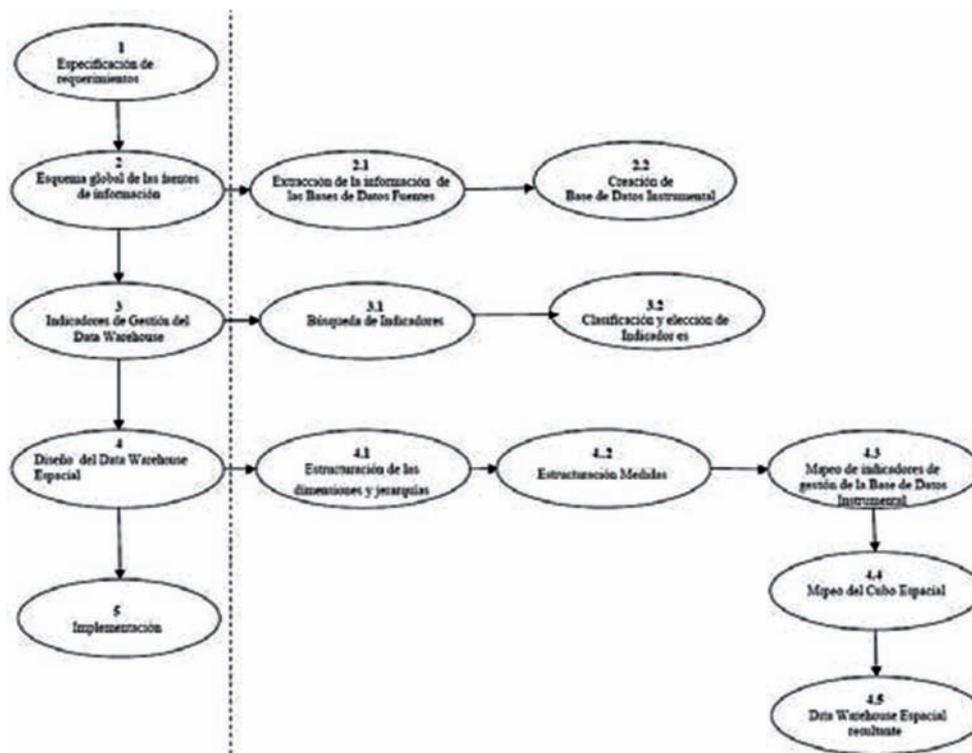
Considerando lo anterior, la solución propuesta es desarrollar un **método** (conjunto de pasos) para crear un Data Warehouse Espacial, el que es aplicado a un caso práctico de un distrito territorial restringido a una comuna, además de presentar la información generada para el análisis en una herramienta de gestión de la información, que facilita la toma de decisiones a organismos o personas encargadas de la administración territorial y de recursos. El trabajo fundamental esta en generar un **esquema global de integración** a partir de bases de datos fuentes de una o varias bases de datos relacionales y espaciales, cada una de ellas con diferente estructura. Para la validación de la propuesta se utiliza un caso de estudio generando un DWE de una comuna de la región Metropolitana como fuente de estudio donde existen una Bases de Datos relacional del último Censo (2003) y una Base de datos de la ficha CAS (2006), además de una Base de Datos Espacial, en todos los casos las bases de datos están en un motor de base de Datos Oracle.

En los apartados siguientes se muestran: Propuesta metodológica del Data Warehouse Espacial DWE, Análisis de los resultados del caso práctico, Conclusiones y Referencias bibliográficas.

## 1. PROPUESTA METODOLÓGICA DEL DATA WAREHOUSE ESPACIAL

La importancia de tener un conjunto de pasos que permitan el diseño e implementación de una base de datos para la gestión de información (DWE), es importante para los diseñadores. Es por ello, que en este trabajo se propone un método que permite implementar un Data Warehouse para consultas de datos espaciales en un distrito territorial de un sector específico. La propuesta se muestra en la figura 1, la que detalla todas las etapas relevantes considerando desde la especificación de requerimientos hasta la implementación y consulta de la información.

FIGURA 1. Propuesta de conjunto de pasos para el diseño Data Warehouse Espacial DWE



A continuación se describirán cada uno de los pasos considerando la implementación de un caso práctico de una base de datos espacial y dos bases de datos relacionales (Ficha CAS y CENSO).

### 1.1 ETAPA 1: ESPECIFICACIÓN DE REQUERIMIENTO

**Descripción:** Esta etapa tiene como finalidad recoger todos los requerimientos de la información de gestión y consultas que desean los usuarios finales. Se debe tener en consideración las fuentes de datos y sus especificaciones para la extracción de la información que permitirán la creación del DWE. Lo complejo de este trabajo, es que los datos están en dos tipos de estructura diferente, espacial como coordenadas y relacional como tablas. Aquí se debe considerar los indicadores de gestión que requiere el usuario final, que serán almacenados en un diseño lógico de estrella, copo de nieve o constelación, generando un diseño de bases de datos multidimensionales.

**Caso Práctico:** Los requerimientos de este proyecto tienen como objetivo o indicador de gestión, determinar el conjunto de personas con mayor riesgo social por áreas de una comuna. La fuente de información se encuentran en una base de datos espacial y dos bases

de datos relacionales implementadas en Oracle. Los indicadores de gestión generales son: 1) Las familias de mayor compromiso social, 2) Índice de distribución geográfica de las familias.

### 1.2 ETAPA 2: ESQUEMA GLOBAL DE LAS FUENTES DE INFORMACIÓN.

**Descripción:** Para la construcción del Data Warehouse Espacial, existen dos alternativas. La primera es poseer una gran instancia de base de datos donde convivan tanto tablas comunes y espaciales de forma tal de integrar toda la información y en su extracción crear un cubo espacial, la segunda es contar con diferentes bases de datos relacionales y bases de datos espacial independientes entre si, se extraen los datos de interés y se crea una tabla de paso o base de datos como esquema global que denominaremos "base de datos Instrumental" y a partir de esta crear los cubos del DWE.

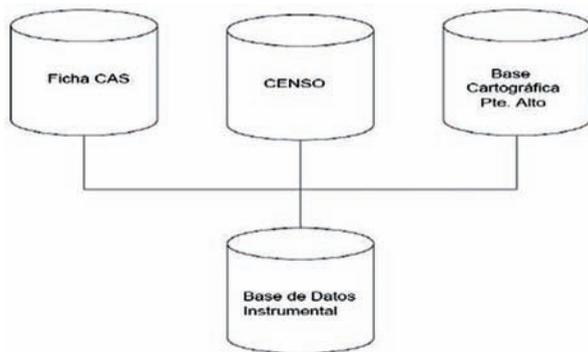
Por tanto, para el paso de la creación de Esquemas global de las fuentes de información (véase figura 1) se optó por la segunda propuesta y se formulan dos subpasos: extracción de la información de las bases de datos fuentes y la creación de la base de datos instrumental, que es la solución que se ha propuesto en este trabajo.

**Caso Práctico:** Las bases de datos fuentes utilizadas son tres, **FICHA CAS**, **CENSO** ambas relacionales y **BASE ESPACIAL**. A partir de estas tres bases de datos se crea un esquema global dando origen a una base instrumental. Existe aquí la primera coyuntura de nuestro caso, que es la importación de los datos que esta en distintos formatos. Para solucionar esta problemática se utilizaron convertidores y un canal de comunicación entre la base de datos Oracle y archivos planos.

Cabe destacar que una vez recolectadas las tablas afines al propósito señalado, se depuran las relaciones y determinan los campos que no aportan al objetivo del cubo espacial, y que debe satisfacer los indicadores de gestión solicitados por el usuario final.

La figura 2 muestra el proceso de obtención del esquema global de las fuentes de información y la interacción entre las bases de datos, que dará origen a la base de datos Instrumental y a partir de esta, se forma el DWE que satisface los requerimientos del usuario final.

**FIGURA 2.** Esquema global de la base de datos instrumental para desarrollar el Data Warehouse Espacial.



En los puntos siguientes se describen las fuentes de datos, sus tablas originales y el diseño de la base de datos instrumental. Una base de datos espacial y dos relacionales.

### 1.2.1 Extracción de la Información Base de Datos Espacial

La Base de datos Espacial de la comuna de Puente Alto es la más importante para la conformación de la base instrumental, ya que sus datos permiten ubicarse geográficamente en el espacio. Aquí se distinguen tres

tipos de datos puntos, líneas y los polígonos los que son clasificados en Hidrografía, Límites, Linderos y Manzanas. El listado de tablas es mostrada en la Tabla 1 y la representación de estas se realiza con un software cartográfico de Oracle (Oracle Spatial Index Avisor), este programa interpreta y dibuja cada uno de los datos de la base de datos espacial.

**TABLA 1.** Base de Datos Espacial Fuente de la Comuna de Puente Alto

HID_ANNOTATION	HIPS_ARC
HID_ARC	HIPS_POINT
HID_LABEL	HIPS_TIC
HID_POLYGON	LABEL_1
HID_TIC	LIM_ANNOTATION
HPS_ANNOTATION	LIM_ARC
LIM_LABEL	LIND_TIC
LIM_POLYGON	MANZANA_ETIQUETA
LIM_TIC	MANZ_ANNOTATION
LIND_ANNOTATION	MANZANA
LIND_ARC	MANZ_TIC
LIND_POINT	MISC_ANNOTATION
MISC_ARC	VER_ANNOTATION
MISC_POINT	VER_ARC
MISC_TIC	VER_LABEL

En esta base de datos espacial, hay un total de 40 tablas, según el análisis desarrollado en este trabajo solo dos de ellas son de interés para conformar la base de datos instrumental, considerando que contienen un campo que es común a las tablas relacionales que son las otras fuentes de extracción (Ficha CAS y CENSO). Los campos en cuestión son "cod\_manzana" que esta incluido en la

tabla MANZANA (dato que forma un polígono) y al dato MANZANA\_ETIQUETA. La figura 3 muestra el esquema para este caso.

Extracción de la información de la base de datos Relacionales: La información que proporciona el organismo interesado para este sistema (MIDEPLAN), se encuentra en una base de datos llamada Ficha CAS, la información contenida en esta base de datos es un requisito imprescindible para la postulación a todo tipo de subsidios que entrega el Estado a las familias en situación de pobreza. El proceso de obtención de los datos de esta base estuvo orientado de acuerdo con los fines del DWE requerido.

La información extraída a un esquema global de tablas en un motor de base de datos Oracle, quedando las que se representan en la Tabla 2.

**TABLA 2.** Tablas de la Base de datos Ficha CAS extraídas

CAS_FAMILIA	CAS_PERSONAS
-------------	--------------

Los datos de la base de datos fuente estaban sin ningún proceso de validación, se realizó una depuración y generaron las tablas CAS\_INGRESO\_PC y CAS\_PUNTAJE. La creación de la tabla CAS\_INGRESO\_PC fue en base a la extracción y depuración de las información de la tabla CAS\_FAMILIA y separar en tramos de ingresos PER\_CÁPITA de a miles de pesos para clasificar a los individuos con mayor riesgo social, el campo RUT del beneficiario fue conservado y eliminados aquellos campos que no tenían ningún rol para el DWE.

La tabla CAS\_PUNTAJE fue creada a partir de la tabla PERSONA\_CAS donde los puntajes fueron extraídos y separados cada 100 puntos (1 al 700), la razón es que las personas que tienen un menor puntaje son más vulnerables, el campo representante fue conservado.

La tabla FAMILIA\_CAS fue reformulada con su información y la de la tabla CAS\_PERSONAS de forma tal que pudiera relacionarse con CAS\_PUNTAJE\_PC cada 10 (1 al 10), CAS\_INGRESO y con la tabla cartográfica MANZANA, como muestra la Figura 3 lado izquierdo.

### 1.2.2 Extracción de la Información de la Base de Datos del CENSO

En la base de datos CENSO están caracterizados todos los habitantes del país por parámetros que reflejan las condiciones de vida. Estos datos están disgregados

por región, provincia y comuna. En nuestro caso, las tablas EDUCACIÓN, PROP\_VIVIENDA, TIPO\_VIVIENDA, PAREJA\_PERSONA son representativas de la condición social de las personas. La recopilación de la información se realiza por un conteo por manzana de los principales parámetros para una estratificación social de las personas que viven en cada sector. La Tabla 3 representa las tablas que conforman la base de datos.

**TABLA 3.** Tablas de la base de datos fuente CENSO

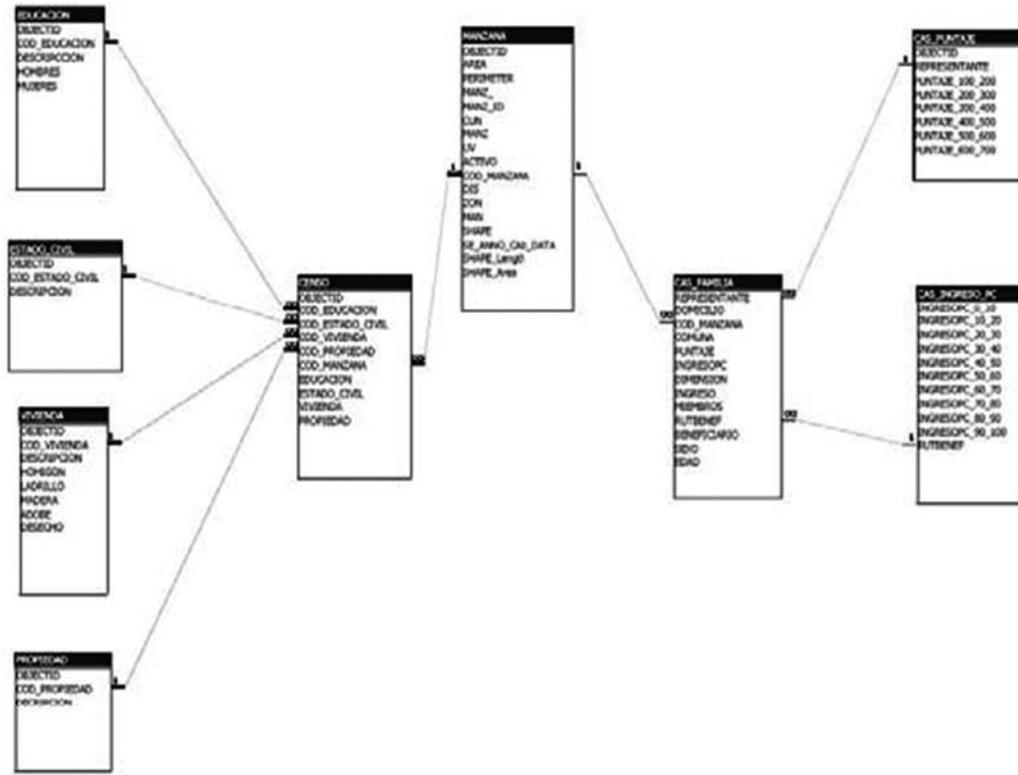
BANO_VIVIENDA	ETNIA_PERSONA
EDUCACION	HOGAR_FUENTE
PISO_VIVIENDA	RELIGION_PERSONA
PROP_VIVIENDA	SEXO_PERSONA
INTEGRANTE_PERSONA	PAREJA_PERSONAS
CONSTRUCCION	PERSONA_FUENTE
TECHO_VIVIENDA	VIVIENDA
TIPO_VIVIENDA	

La extracción de la información de la base instrumental comenzó con la selección de las tablas que estaban en un formato distinto de base de datos, por lo que se realizó un proceso de transformación de la data al formato requerido, se eliminaron los campos que no tenían relación con el trabajo, y se establecieron las relaciones correspondientes, la base de datos resultante se muestra en la figura 3 lado derecho. La creación de una nueva llave en cada una de las tablas son: cod\_estructura, cod\_persona, cod\_propiedad y relacionar a su vez la tabla CENSO\_COMUNA con la tabla espacial MANZANA por el cod\_manzana.

### 1.2.3 Base de datos instrumental

Esta Base de Datos es el elemento fundamental para la creación del Data Warehouse Espacial, aquí está el esquema global depurado de las tres tablas mencionadas anteriormente. La figura 3 muestra el esquema resultante de la base de datos instrumental que es la unión la extracción de ambas bases de datos fuentes.

FIGURA 3. Base de Datos Instrumental



La construcción del Data Warehouse Espacial resultante de la base de datos instrumental, tiene dos posibilidades para su desarrollo: la primera crear un diseño estrella para el Data Warehouse y la segunda alternativa es crear un diseño constelación. Para esto, se realizó un análisis de la estructura y coherencia de los datos y se diseña un cubo para satisfacer los indicadores de gestión de la base de datos CENSO y otro cubo para los indicadores de gestión de la base de datos Ficha\_CAS, generando un diseño lógico de estrella para cada cubo y así se implementan las metas propuestas en el Data Warehouse Espacial. Otro análisis, es establecer jerarquías de las dimensiones, para que puedan utilizar algunas operaciones Roll Up y navegar en ellas con un proceso de agregación.

### 1.3 ETAPA 3: INDICADORES DE GESTIÓN DEL DWE

**Descripción:** En este paso se debe complementar los elementos de gestión para los cuales se construye el DWE, que debe ser concordante con la información extraída en la etapa anterior. En nuestro trabajo se decide realizarlo en una etapa posterior a la extracción

de datos, para estar seguros que la información que se requiere para cada indicador de gestión existe en la base de datos instrumental.

Considere que los indicadores de gestión son aquellos que suministrarán una información medible proporcionada por los Data Warehouse Espacial. Su característica principal es que son complementarios, es decir ambos proporcionan datos acerca de un sólo elemento que interesa analizar y medir, a partir de datos ya existentes por un tiempo. Esta etapa está compuesta por dos subetapas: Búsqueda de los indicadores y clasificación y selección de indicadores.

**Caso Práctico:** En nuestro caso de estudio, uno de los indicadores de gestión que se quiere implementar en un DWE, que permita obtener una información de análisis con respecto a -las manzanas de la comuna de Puente Alto-, que es el elemento cartográfico donde los organismos o personas que deseen tomar decisiones respecto a -la distribución de recursos-, para ello se requiere de una herramienta que entregue esta información. En los puntos siguientes serán analizados estos indicadores con mayor detalle.

### 1.3.1 Indicador 1: Las familias de mayor compromiso social

El indicador de la de la base de datos fuente Ficha\_CAS, es el puntaje otorgado a las familias -mientras menor sea este mayor es su compromiso social- hasta llegar a la extrema pobreza. Las familias que postulan a la ayuda gubernamental ingresan a un proceso de entrevistas con asistentes sociales que registran sus apreciaciones en base a una serie de parámetros que no se analizan en este trabajo, si destacaremos uno de ellos que es el ingreso mensual, la cual es una fuente muy importante de medición.

Para la base de datos fuente CENSO existen múltiples indicadores que reflejan el -estándar social de las familias de la comuna-, En este trabajo se han elegidos los que tienen mayor relevancia como el tipo de vivienda (que se describe desde vivienda una sólida hasta una creada con desperdicios), propiedad (que muestra los propietarios, arrendatarios, viviendas gratuitas o cedidas por el trabajo), la educación (desde Universitaria hasta sin educación) y por último estado civil (que mide la constitución del vínculo entre parejas desde casados hasta viudos).

### 1.3.2 Indicador 2: Índice de distribución geográfica de las familias

Este índice no se analizará en este trabajo, pero se tiene claro que puede ser de gran ayuda para los planificadores, ya que se obtiene, la cantidad de habitantes por manzana. Este indicador es medible fácilmente, ya que en cada una de las manzanas esta su área en metros cuadrados y de esa forma se consigue un índice de densidad poblacional.

## 1.4 ETAPA 4: DISEÑO DEL DATA WAREHOUSE ESPACIAL

*Descripción:* Esta etapa tienen como objetivo generar los cubos y diseño lógico que permitan implementar los indicadores de gestión en las tablas necesarias en un diseño multidimensional estrella, copo de nieve y otro que sea requerido, a partir de los datos de la bases de datos instrumental.

Este paso tiene como subpaso: Estructura de las dimensiones y jerarquías, estructuración de las medidas, mapeo de los indicadores de gestión desde la base de datos instrumental, para la creación de cubos espaciales para obtener el diseño lógico del DWE.

*Caso Práctico:* El primer cubo se obtiene, según la especificación de requerimiento "Cantidad de Familias con Marginación Urbana por Manzana". Se escoge de la

base instrumental las tablas y los datos obtenidos de la base de datos Ficha\_CAS que reflejan las familias que tienen problemas de extrema pobreza por manzana.

Las tablas que fueron seleccionadas son: Ingreso Per\_cápita, Puntaje\_del\_Individuo, Manzana a partir de ellas se crean las **medidas** siguientes:

**Ingreso:** El número de personas que perciben ingresos en los distintos rangos definidos por manzana.

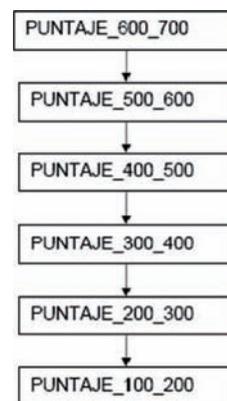
**Puntaje:** El número de familias que están en los rangos dispuestos por manzana

La figura 5 parte superior, muestra el cubo que reflejará la conformación del DWE. Dado el análisis anterior los datos de las tablas que generarán las dimensiones son: Ingreso\_per\_capita (se extraen datos desde la tabla cas\_ingreso\_pc), Puntaje que obtiene los datos desde la tabla cas\_puntaje y manzana (que obtiene datos desde la tabla de datos espaciales).

La **Tabla de Hechos** en un diseño lógico de estrella, mostrado en la figura 5 parte inferior, donde están las tablas y los datos que permiten contar los puntajes de ingreso\_per\_cápita como medida que permitirán crear el DWE y reflejar por manzana los problemas de marginación.

En un análisis de las dimensiones de la tabla Cas\_familia extraída de la base de datos fuentes Ficha\_CAS, se encontró sólo un campo que indicaba el puntaje de los postulantes para poder definir la dimensión, por lo que este dato se clasificó de 100 puntos, de tal forma que los habitantes puedan ser encasillados en cualquiera de estos tramos. La figura 4 muestra un ejemplo de esta dimensión en jerarquías.

FIGURA 4. Jerarquía de los Puntajes



El mismo análisis se realizó para el atributo Ingreso\_familiares de la misma dimensión, donde la jerarquía esta jerarquizada en forma de que los ingresos están acotados de a 10 mil pesos. Para cada uno de estos casos se desarrollo la operación agregación respectiva, que permite el escalafón de los ingresos que se clasificaron de acuerdo a esta medida. del mismo modo la tabla Manzana esta estructurada con dos elemento de jerarquía: área en metros cuadrados y perimetro en metros.

#### 1.4.1 Mapeo de indicadores de gestión de la Base de Datos Instrumental

En esta etapa se desarrollan dos cubos una para el indicador que se obtienen de la Ficha\_CAS y el otro cubo que se obtienen de Iso indicadores del CENSO.

#### 1.4.2 Mapeo de Cubo para la Ficha CAS

El mapeo de fragmentos es una función que muestra cómo se corresponden los objetos del modelo conceptual (figura 5) a un diseño lógico (Peralta, 2001). Realizando un seguimiento de la función de mapeo, se seleccionan 3 tablas de la base de datos instrumental que representan

las dimensiones de Ingreso\_per\_cápita, la tabla geométrica manzana y la tabla de hechos CAS\_FAMILIA.

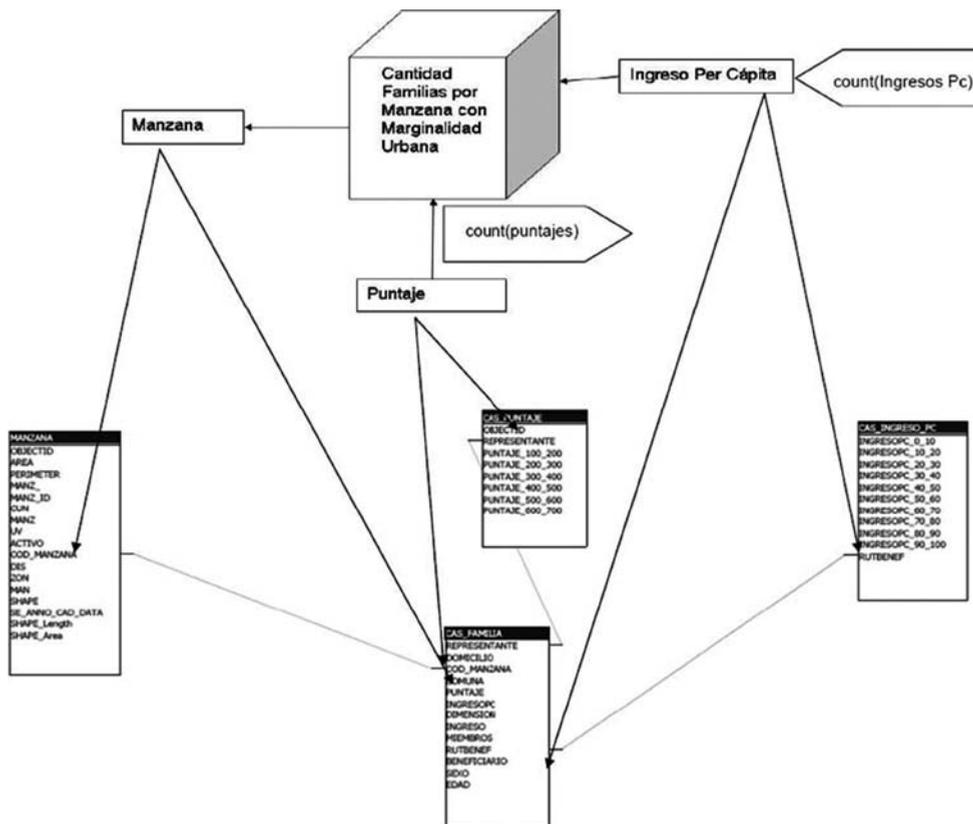
La consulta se muestra a continuación:

```
SELECT al1.rutbenef, al2.domicilio, al1.ingresopc_0_10,
al1.edad, al3.cod_manzana FROM cas_ingreso_pc al1,
cas_familia al2, manzana al3 WHERE (al2.rutbenef=al1.
rutbenef and al3.cod_manzana=al2.cod_manzana) and
(al1.rutbenef=14134476)
```

Esta misma consulta es desplegada en un software cartográfico como Arcgis, se puede apreciar la diferencia con respecto a una base tradicional, dado que la base instrumental es de tipo espacial donde se aprecia el plano urbanístico de la comuna y en especial la manzana seleccionada por la consulta (Medina, 2009).

Los mapeos de las funciones realizadas a las tablas de la base de datos instrumental de puntajes e Ingresos\_per\_cápita, muestran que las funciones definidas en la vistas SQL coinciden con el fragmento del cubo que quiere mapear. La figura 5 muestra el mapeo de cubos desde la base de datos instrumental.

FIGURA 5. Mapeo cubo de la ficha CAS

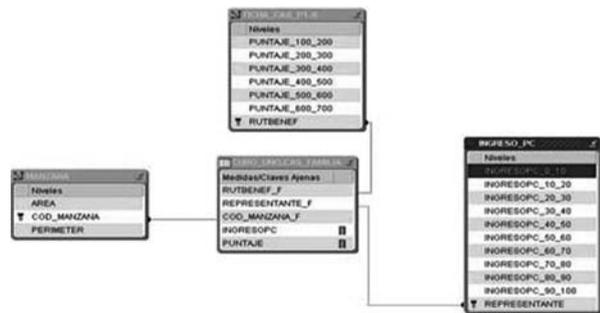


La funciones definida es contar las familias por manzana que tengan puntajes de Ficha\_CAS y además de Ingreso\_per\_cápita en un rango determinado. En este caso será con la instrucción SQL count (puntaje) y count (Ingreso\_per\_capita).

## 2. Data Warehouse Espacial Resultante de la Base de Datos Fuente Ficha\_Cas.

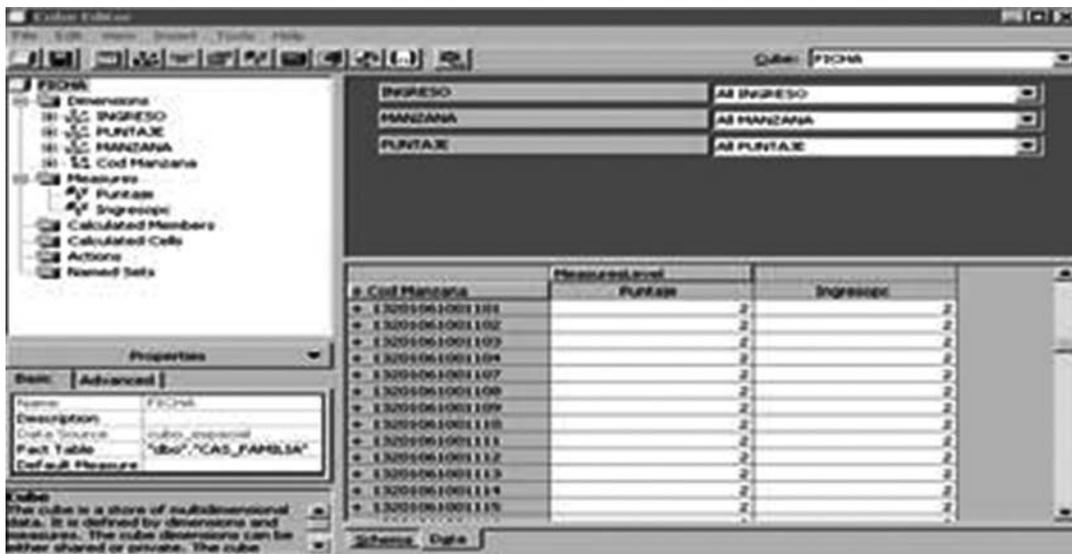
Después de realizar todas las etapas y subetapas anteriores, se genera el cubo como diseño lógico de estrella representado en la figura 6. Cabe señalar que la tabla manzana es la que contiene el campo espacial y representa un polígono como se ha apreciado anteriormente. Los resultados son mostrados en Vistas Materializadas (SnapShot Long) de la Figura 7.

FIGURA 6. Data WareHouse Espacial



La figura 7 muestra una vista del cubo desarrollado consultado en SQLSERVER 2000, los datos cartográficos en este caso son representados como binario largo.

FIGURA 7. Resultados Cubo Ficha CAS



Data Warehouse Espacial de la base de datos fuente CENSO

En el segundo cubo se profundizará el análisis de las familias desposeídas por manzana con los datos de la base de datos fuente CENSO. Primero se definen las dimensiones y medidas, luego el mapeo del cubo desde los datos de la base de datos instrumental para finalmente crear el diseño lógico de estrella.

Las dimensiones elegidas para formar el DWE del caso de estudio son: Vivienda, Educación, Manzana, Estado Civil, Propiedad.

Las **medidas** están definidas de la siguiente forma:

**Vivienda:** El número de casas, departamentos, piezas, medias aguas por manzana.

**Propiedad Construcción:** El número de propietarios, arrendatarios, con casa gratuita y cedida por el trabajo.

**Educación:** El número de Universitarios, alumnos enseñanza media, básica y sin educación.

**Estado civil:** El número de casados, separador, viudos.

Según esto, las familias por manzana que tienen mayor índice de pobreza son los datos que clasificados como - el peor tipo de vivienda, educación y techo- ya que indican mayor precariedad desde el punto de vista social.

El índice de familias más desposeídas por manzana estará compuesto por -el número de personas con las peores condiciones- de acuerdo con las medidas que el cubo espacial presente después del proceso.

De acuerdo con la base de datos instrumental, para el cubo la tabla de hecho son extraídas de la base de datos fuente Censo y las dimensiones son: Educación, Vivienda, Construcción y Manzana y Estado Civil.

Para cada dimensión se analizan las jerarquías correspondientes para poder navegar por Operaciones de Agregación y con consultas Roll Up. En las dimensiones propiedad y estado\_civil no es posible establecer los diagramas de jerarquía de dimensión, en cambio para la dimensión educación, se pudo crear la jerarquía sexo. En el caso de Manzana, mal igual al caso anterior, las jerarquías son perímetro y área. Para el caso de la dimensión infraestructura se han elegido las jerarquías de dimensión más representativa como: Hormigón, Ladrillo, Madera, Adobe y Desecho

## 2.1 Mapeo de Cubo para la base de datos fuente CENSOS

La función mapeo definida será contar el número de personas por nivel educacional por manzana. Para el caso de este seguimiento de fragmentos se puede

hacer un de los registros desde la tabla Educación a la de Manzana y la consulta es -cuantas personas en la manzana (cod\_manzana) 13201011001009 tienen estudios en Humanidades-.

La consulta con los parámetros antes mencionados sería la siguiente:

```
SELECT al3.cod_educacion, al3.descripcion, al1.educacion, al2.cod_manzana FROM censo al1, manzana al2, educacion al3 WHERE (al2.cod_manzana=al1.cod_manzana and al3.cod_educacion=al1.cod_educacion) and ((al3.cod_educacion=60 and al2.cod_manzana=13201011001009))
```

```

Educacion      7
Cod Educacion  60
Descripcion    HUMANIDADES
Cod Manzana1  3.201.011.001.009
  
```

La consulta realizada a nivel de base de datos espaciales el resultado se reflejará en la figura 8

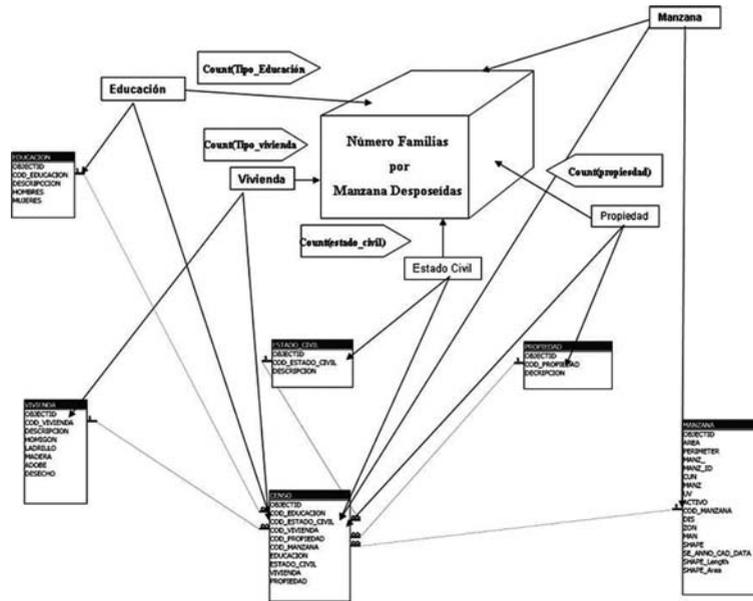
FIGURA 8. Consulta de Educación en Base Espacial



En la figura 9 esta el mapeo de cubo correspondiente, la función de mapeo es contar en educación el número de personas por nivel de enseñanza, en vivienda es el número de construcciones por su tipo, para propiedad es el número

de persona dueñas, arrendatarias etc. y finalmente en estado civil contar el número de personas casadas, separadas y todos estos totales son por manzana.

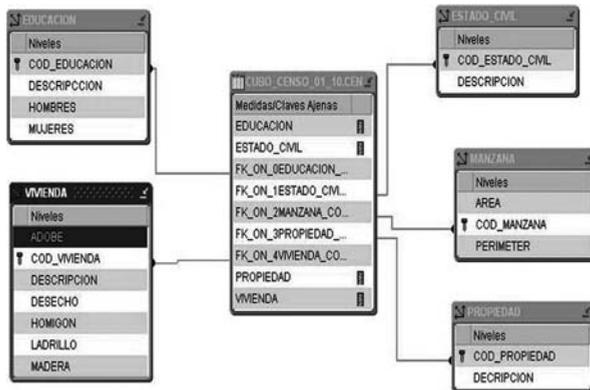
FIGURA 9. Mapeo del cubo espacial del CENSO



## 2.2 Data Warehouse Espacial Resultante de la Base de Datos Fuente CENSO

Finalmente después de realizar todos los pasos de análisis y diseño, es implementado el cubo en un diseño estrella tal como lo representada en la figura 10. Hay que señalar que la tabla manzana es la que contiene el campo espacial, que realiza el nexo entre las dos bases de datos. Los resultados son mostrados en Vistas Materializadas(SnapsShot Long) de la figura 11.

FIGURA 10. Data WareHouse Espacial Censo



Para poder visualizar el cubo y hacer un análisis del indicador de gestión, el cubo es transportado al Análisis Manager de SQL Server, aquí como consulta, son obtenidas desde cada una de las medidas por manzana de la comuna en cuestión. Una consulta de este cubo lo muestra la Figura 11.

FIGURA 11. Resultados Cubo Censo de la Comuna

Cod Manzana	Educacion	Estado Civil	Vivienda	Propiedad
1.000001.10001001	10,190	10,190	10,190	10,190
1.000001.10001002				
1.000001.10001003				
1.000001.10001004				
1.000001.10001005				
1.000001.10001006				
1.000001.10001007				
1.000001.10001008				
1.000001.10001009				
1.000001.10001010				
1.000001.10001011				
1.000001.10001012				
1.000001.10001013				
1.000001.10001014				

### Paso 5: Implementación

La implementación de los cubos, en este caso fue realizada en SQL Server, y los resultados son mostrados en las figuras 7 y 11.

## 3. ANÁLISIS DE LOS RESULTADOS DEL CASO PRÁCTICO

Un administrador o planificador, con la información proporcionada por el cubo espacial, permite hacer una clasificación de cada una de las manzanas que integran la comuna, de forma tal de diferenciar cuáles son las áreas con mayor compromiso social.

En el caso ejemplo, las manzanas pueden clasificarse individualmente de acuerdo con las medidas definidas en los DWE de cada uno de los cubos obtenidos de las bases de datos fuentes Ficha\_CAS y CENSO, extraídos los atributos: puntaje, ingreso\_per\_cápita, educación, propiedad de la vivienda, tipo vivienda, y conformación

del núcleo familiar, además de vistas en el mapa de la comuna. Los indicadores de gestión obtenidos en este caso de estudio, son un apoyo muy importante a los encargados de administrar y planificar la entrega de servicios y recursos a un área geográfica determinada, como una manzana, y así eliminar la ceguera territorial en que ellos están sumidos, por la falta de Software que les permitan gestionar la información con la que cuentan.

El método aplicado para la creación de un DWE, puede ser usado a niveles territoriales aún mayores como una región o la totalidad del país, la problemática es poder obtener información de los distintos organismos públicos para la construcción del cubo.

Un aspecto significativo para los planificadores en una comuna son los avances de sus habitantes, esto puede medirse en base a la información entregada por el DWE extraídos en períodos de tiempo determinados. Este desarrollo incremental del DWE es posible ya que se puede anexar más información a la base instrumental y realizar el proceso nuevamente.

El caso práctico nos muestra la importancia de un DWE para la gestión de un área determinada, en el ejemplo son descritos los principales puntos de análisis que permiten al usuario final contar con una información mas acorde a la temática que ellos manejan.

El Data Warehouse Espacial creado a partir de las bases de datos fuente Ficha\_CAS y Data Warehouse Espacial creado a partir de las bases de datos fuente CENSO, crean dos cubos independientes implementados en el DWE y cada uno de ellos poseen la información descrita de cada una de las manzanas de la comuna, permitiendo visualizar las condiciones de la población con mayor información de gestión en comparación a los métodos utilizados actualmente. Esta información puede desplegarse con software cartográficos, los que tienen asociadas herramientas PL/SQL que pueden realizar consultas más especializadas, mostrando sus resultados en el mapa cartográfico de las comunas con datos existente y calculados al interior del Data Warehouse Espacial.

Finalmente, este trabajo tienen como objetivo mostrar los pasos con los cuales se puede desarrollar una nueva innovación tecnológica para la gestión territorial, ahora es trabajo de autoridades y jefes de servicio o personas con iniciativas similares que aprovechen esta método para mejorar las condiciones de los habitantes de su comuna.

#### 4. CONCLUSIONES

Data Warehouse Espacial como se ha apreciado es un instrumento que proporciona información que mejora estos procesos de gestión, ya que son incluidos factores

de ubicación espacial, datos estadísticos, datos reales todos unidos en sólo un conjunto de antecedentes para la persona y organismo encargado de gestionar los recursos tanto gubernamentales como privados.

Entregar un conjunto de pasos para la creación de un Data Warehouse Espacial descrita en este trabajo servirá para fomentar a las instituciones desarrollar proyectos de similares características al expuesto. El contar con una explicación detallada de las fases de construcción contribuye a profundizar el entendimiento a las personas que tendrán a cargo esta iniciativa, los administradores y planificadores.

En cuanto a la identificación del problema: El área pública es uno de los principales propulsores de innovaciones que apuntan a mejorar la gestión en todos los ámbitos, en especial en el área de la información para la toma de decisiones, pero existen algunas limitaciones, no todas las reparticiones conocen las implementaciones tecnológicas que otras han logrado, por lo que no se han aprovechado de éstas innovaciones que perfeccionarían muchos de sus procesos internos.

En este caso el Data Warehouse y las Bases de Datos Espaciales esta última ampliamente conocidas y utilizadas por algunos organismos públicos. Es así que este proyecto une estas dos materias con el fin de mejorar un aspecto fundamental, mejorar los procesos de gestión por medio de un método para la creación de un Data Warehouse Espacial.

Recolección de información, análisis de los datos: En la recolección de la información existieron dos problemas que dificultaron la concreción de esta iniciativa, la primera, es la excesiva reserva de los datos y la segunda, es que la información esta orientada exclusivamente a aplicaciones estáticas, y por consiguiente no existe una integración de los datos en las bases estudiadas. Estos datos eran de mala calidad e incompletos por lo que se realizó un proceso de racionalización de la información previo estudio de los diseños de las bases de datos, para estructurar sus tablas, modificando y eliminando campos para construir nuevas relaciones, traspasando esta nueva organización a la base de datos instrumental.

Arquitectura de la Base de Datos Instrumental: La creación de la base de datos instrumental a partir de un grupo global de bases de datos relacionales y espaciales es el resultado del proceso de recolección, clasificación y análisis de datos.

Tipos de datos que conforman la Base de Datos Instrumental: Fueron utilizados dos tipos de datos en las tablas de la Base instrumental, el primer tipo son

los datos clásicos de un base de datos relacional y el segundo es el dato de coordenadas de una base de datos espacial, en su conjunto conforman las bases de datos de objeto relacional espacial, por el carácter de este tipo de datos es muy difícil comprenderlo.

Generación de los Indicadores: Los indicadores de gestión para el DWE son creados con una finalidad específica de poder analizar y medir características específicas de datos espaciales y clásicos. Para el caso específico de este estudio se buscó medir el número de personas con mayor compromiso social en una manzana, siendo este último el dato espacial en particular un polígono. Esto significa que los indicadores medirán los datos que intrínsecamente están relacionados al elemento cartográfico de forma tal que puedan ser representados en una figura bidimensional de una manzana.

Conjunto de pasos para la construcción del Data Warehouse Espacial: Es caracterizada por un método presentado, especificando los requerimientos, recolectando información de las bases de datos fuentes para generar la Base de Datos Instrumental. El paso siguiente es determinar los indicadores de gestión para identificar las dimensiones y medidas para finalmente obtener el diseño del DWE. El caso práctico de este proyecto es crear dos cubos a partir de las fuentes de datos Ficha\_CAS y CENSO utilizando para ambos el mismo método.

La conformación de más de un cubo a partir de una base instrumental ocurre muy a menudo, ya que no existe compatibilidad entre las tablas, pero es destacable el hecho de que pueden ajustarse perfectamente y reunir sus resultados cuando existe un dato en común en este caso desde la tabla Manzana.

Actualizaciones y mantención: Hay que destacar que una de las características más importantes de los Data Warehouse Espaciales, es que son escalables, es decir puede anexarse más información en el tiempo, y también son flexibles, es decir, pueden ser modificados para generar cubos con otros objetivos de estudio. Este es un aspecto donde los Data Warehouse Espaciales y no pierde vigencia, al contrario es una invaluable herramienta para establecer los progresos o retrocesos de la población en cualquier área urbana o rural.

## 5. TRABAJOS FUTUROS

Desarrollar este método en otros motores de bases de datos, como PostgreSQL que incluye el dato espacial y asociarlo a software cartográficos libres (PostGis), que despliegan esto registros cartográficos de igual forma

de aquellos debe pagarse licencias. Esto abaratará los costos y permitirá una masificación de proyectos similares tanto a nivel comunal como regional.

Incluir en esta metodología en otro tipo de elementos que soportan los motores de base de datos como imágenes, textuales, sonido etc. que contribuirán a enriquecer este tipo de Data Warehouse Espacial.

## 6. REFERENCIAS

- [1] Jan Van Berkel (1997): Documento en Línea. Datawarehouse. Where to locate Gis, Logisterion Automatisering
- [2] Terri L. Cohen, Julie R. Baitty (2005): Health Resources and Services Administration Geospatial Data Warehouse. Integrating Spatial and Tabular Extract, Transform, and Load processes. <http://gis.esri.com/library/userconf/proc05/papers/pap2405.pdf>
- [3] Michalis Petrakos (2002): Design of Geographical and Statistical Database. Information Society Technologies – Statistical Atlas of the European Unión. [http://www.statlas.org/project/deliverables/wp2/STL\\_GSDB\\_D2.2.pdf](http://www.statlas.org/project/deliverables/wp2/STL_GSDB_D2.2.pdf)
- [4] ESRI (1998):.ESRI Shapefile Technical Description .An ESRI White Paper July 1998. <http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf>
- [5] Douglas D. Nebert (1994): GeoSpatial Data Clearinghouse Systems Design Workshop Summary Report, Pacific Meridian <http://gis.esri.com/library/userconf/proc95/to200/p196.html>
- [6] Joaquín A., Rodríguez (2000): Minería de Datos y Sistemas de Información Geográfica. [http://www.mappinginteractivo.com/plantilla-ante.asp?id\\_articulo=142](http://www.mappinginteractivo.com/plantilla-ante.asp?id_articulo=142)
- [7] Reinaldo Medina (2009): Propuesta de un Data Warehouse Espacial para la Gestión de Distritos Territoriales. Tesis de magister en tecnologías de la información. Universidad Santiago de Chile.
- [8] Verónica Peralta (2001): Diseño Lógico de Data Warehouse a partir de Esquemas Conceptuales Multidimensionales, Tesis de Maestría, Universidad de la Republica, Montevideo, Uruguay.