

Desarrollo de una solución de inteligencia de negocios para integrar datos de censo y encuesta en un SIG

R. Cornejo, M. Navarrete,
R. Valdivia, P. Aroca, S. Aracena

Resumen

En este trabajo se presenta el desarrollo de una solución de Inteligencia de Negocios para gestionar el gran volumen de datos generados a partir de dos instrumentos de recolección, como lo son el censo nacional de personas, hogares y viviendas y la encuesta de hogares (CASEN); con el objetivo de desarrollar una plataforma de gestión de bases de datos en variables socioeconómicas para consulta y visualización cartográfica micro territorial integrada. Esta solución involucra la implementación de un almacén de datos intermedio que permite tener una visión conjunta de ambas fuentes de datos. La integración de datos se realiza mediante la técnica de matching espacial que requiere unas condiciones de homogeneidad entre ambos instrumentos respecto de las características de la muestra, de las unidades administrativas de desagregación espacial y respecto de las preguntas de recolección de datos. Utilizando los códigos de localización espacial incorporados en una base de datos espacial y a través de un Sistema de Información Geográfica (SIG); se añaden herramientas de visualización cartográfica, lo que facilita la observación y análisis de las relaciones espaciales entre las unidades geográficas administrativas, así como la observación y análisis de las particularidades intrarregionales. Como aplicación, se presenta la distribución del ingreso per cápita de los hogares en unidades intracomunales, información que no se ha obtenido hasta ahora puesto que los datos de ingreso de los hogares que entrega la encuesta de hogares no permiten una desagregación menor a la comuna.

Palabras clave: Inteligencia de negocios. Matching espacial. Base de datos espacial. Sistema de Información Geográfica.

Abstract

In this paper we present the development of a Business Intelligence solution to manage the large volume of data generated from two data collection instruments, such as the National Census of people, households and housing and the Household Survey with the aim to develop a platform management integrated of data-

ses for consultation of socioeconomic variables microterritorial with cartographic visualization. This solution involves the implementation of an Intermediate Data Warehouse that allows an overview of both data sources. Data integration is done by Spatial Matching technique that requires homogeneity conditions between the two instruments on the sample characteristics, the administrative units of spatial disaggregation and the answer or questions in the data. Using spatial location codes incorporated in a Spatial Database a through Geographic Information System (GIS), are added cartographic visualization tools, which facilitates the observation and analysis of spatial relationships between administrative geographical units and as observation and analysis of intra-regional particularities. In the application, we present the distribution of income per capita household in intracomunal units, information that has not been achieved so far as the data of household income in the household survey does not allow one disaggregation less than the Commune.

Keywords: Spatial Matching. Spatial Data Base. Business Intelligence Solution. Geographic Information System.

R. Cornejo (rcornejo@uta.cl), Universidad de Tarapacá. **M. Navarrete** (monica.navarrete@gmail.com), Universidad de Tarapacá. **R. Valdivia** (rvaldivi@uta.cl), Universidad de Tarapacá. **P. Aroca** (patricio.aroca@uai.cl), Universidad Adolfo Ibáñez. **S. Aracena** (sebastian.aracena.aguirre@gmail.com), Universidad de Tarapacá.

INTRODUCCIÓN

Sin duda que la contribución de los Sistemas de Información Geográfica (SIG) al tratamiento y análisis de los datos de tipo espacial han facilitado el estudio de eventos socioeconómicos claves para la elaboración de políticas públicas relacionadas con, por ejemplo, programas de superación de la pobreza, programas de control ambiental y/o programas de intervención urbana. Uno de estos eventos ha sido la representación espacial de la distribución del ingreso per cápita de los hogares (véase Elbers, Lanjouw y Lanjouw, 2003; Minot y Baulch, 2005) utilizado en la construcción de mapas de pobreza regional aplicado en más de 40 países para obtener distribuciones espaciales de la pobreza, riqueza y/o desigualdad y que como se reconoce en Elbers y Lanjouw (2012), buscan orientar espacialmente la intervención pública ayudando así a comprender los beneficios de la descentralización con un mayor foco en el desarrollo centrado en la comunidad. Para Chile, aplicaciones de lo anterior lo vemos en Agostini, Brown y Góngora (2008), en Agostini et al (2010) y en Modrego et al (2008, 2012). Sin embargo, en la construcción de estos mapas de pobreza, los SIG son subutilizados al asignarles sólo la representación visual del ingreso o tasa de pobreza geográfica. En este trabajo se propone el desarrollo de una solución de inteligencia de negocios para integrar y gestionar datos de encuesta y censo en un SIG, logrando de esta manera no sólo generar los mapas de pobreza sino que aprovechar la información contenida en la base de datos espacial para la observación de las interrelaciones espaciales del efecto vecindad en los niveles de pobreza y/o riqueza en niveles incluso micro territoriales. Esta solución contempla la extrapolación territorial de datos a través de la correspondencia entre códigos de ubicación territorial que surge como resultado del uso de la tecnología ETL (Extract, Transform and Load) y un proceso de ingeniería inversa. La base de datos contiene información integrada, indicadores generados por extrapolación espacial (en niveles de desagregación que esta no provee y que se obtienen mediante un matching espacial) y una plataforma de visualización cartográfica a través de un servidor de mapas; con lo cual se espera mostrar la aplicación y utilidad de los SIG en la observación, desarrollo e integración de información socioeconómica de los hogares. Por tanto; en los siguientes apartados mostramos la creación de un repositorio de datos espacial, la implementación de una herramienta de escritorio para el manejo de sistemas de información geográfica y la creación de una plataforma de visualización cartográfica a través del levantamiento de un servidor de mapas. En el apartado de metodología, se exponen las herramientas tecnológicas utilizadas y como aplicación, presentamos la cartografía micro territorial del ingreso per cápita de los hogares en las regiones XIII, VI y VII de Chile.

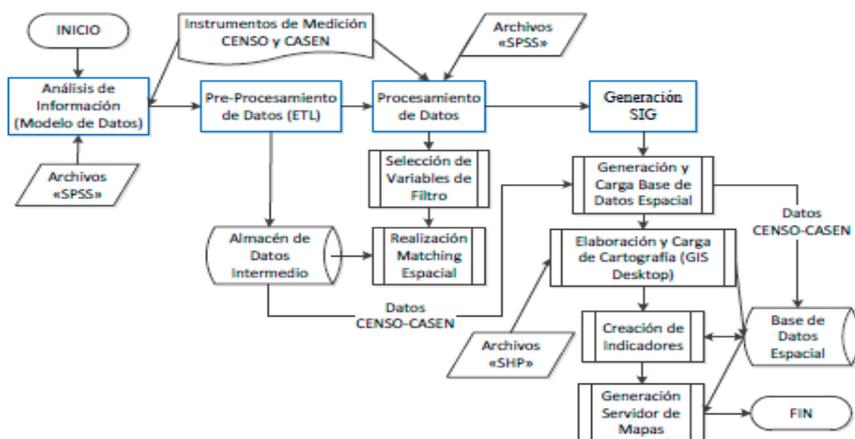
METODOLOGÍA

La solución propuesta utiliza tecnología ETL, perteneciente a la inteligencia de negocios y los SIG. El ETL es una tecnología de integración de datos que se utiliza en proyectos de implantación de inteligencia de negocios, el cual permite extraer datos alojados en diversas fuentes de información, transformarlos según las necesidades del negocio y cargar estos en los entornos de destino. Inteligencia de negocio (R. Kimbal, 1998) constituye un enfoque estratégico para orientar sistemáticamente el seguimiento, la comunicación y la transformación relacionada al débil conocimiento de la información procesable en la cual se basa la toma de decisiones, siendo una de sus tareas más significativas; la del diseño y construcción de los almacenes de datos o data warehouse (DW) conocidos como “una colección de datos orientados a un ámbito (empresa, organización), integrada, no volátil y variante en el tiempo; que ayuda al proceso de los sistemas de soporte de decisiones” (W. H Immon. 2002), mientras que un SIG brinda una representación gráfica de la información geográficamente referenciada. Un SIG se basa en principios formales de matemáticas discretas, modelos de datos y geometría computacional; su desarrollo, involucra la aplicación de nuevas tecnologías de la información: estándares e ingeniería de software, almacenes de datos, web- SIG, metadatos, ambientes y lenguajes visuales, entre otros (P. Lejia Luna, 2010). La mayor utilidad de un sistema de información geográfico está intrínsecamente relacionada con la capacidad que éste posee de construir modelos o representaciones del mundo real a partir de las bases de datos digitales, esto se logra aplicando una serie de procedimientos específicos que generan información para el análisis (GEOINFO, 2012).

El software utilizado es el gvSIG (N. Sillero P. Tarraso, 2010) que es una de las herramientas de código abierto, orientada al manejo de la información geográfica, más potentes del mercado. Además de ser una solución para la interpretación de datos en tecnología SIG, gvSIG se utiliza para un proceso adicional de extracción, depuración y carga, que permite el traspaso de archivos Shape (que contienen la cartografía del país) a datos espaciales georreferenciales en PostgreSQL (PostgreSQL, 2012), funcionalidad que se logra al añadir el módulo PostGIS, que permite dar soporte a objetos geográficos transformando una base de datos común en una base de datos espacial que puede ser utilizada en SIG y servidores de mapas (OSGeo, 2012). Por último, esta propuesta plantea la creación de una aplicación Web para la publicación de datos espaciales, mediante la tecnología Internet Map Server (IMS, Servidor de Cartografía Digital), con el fin de permitir a los usuarios visualizar, consultar y analizar información geográfica a través de la red, tanto en modo vectorial como ráster, realizado con un servidor de mapas denominado MapServer con el Framework Pmapper. La Figura 1 muestra el diagrama de los

procesos y sub-procesos utilizados para cada caso en particular. El análisis de información o estudio de las fuentes de información (CENSO 2002 y CASEN 2003) necesarios para elaborar los modelos relacionales a utilizar. El pre-procesamiento de datos, donde se lleva a cabo el proceso ETL en un almacén de datos intermedio (ADI). El procesamiento de datos donde se implementa el matching espacial entre los instrumentos de medición CENSO y CASEN y por último la generación SIG en el que se construye la plataforma de visualización cartográfica a través de diversos sub-procesos y herramientas de libre disposición.

Figura 1: Diagrama de procesos y sub-procesos realizados en el desarrollo de la tesis.

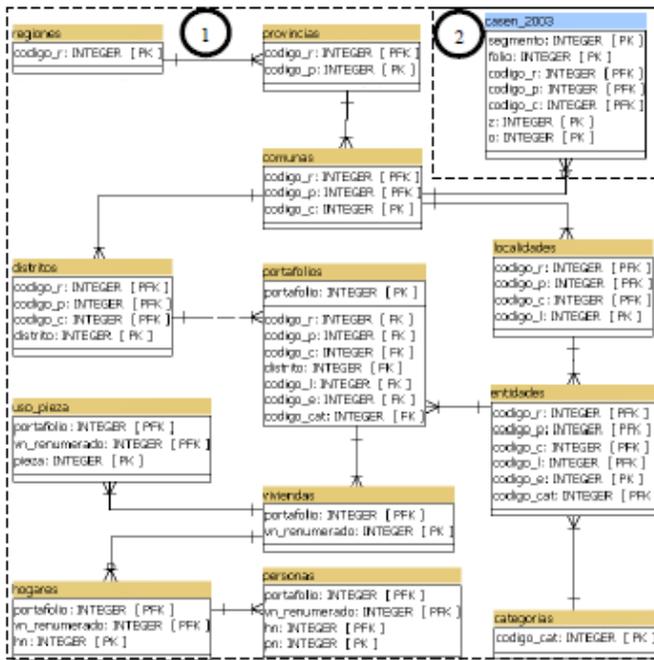


Fuente: elaboración propia.

INTEGRACIÓN DE DATOS MEDIANTE MATCHING ESPACIAL

La integración de datos entre encuesta y CENSO utilizó un proceso de ingeniería inversa que permitió reconstruir el modelo de datos de CENSO mediante el estudio de la semántica de la información contenida en la encuesta respectiva. El modelo relacional obtenido "CENSO-CASEN", fusiona la información de ambos instrumentos en un único modelo, de modo que se aprovecha la cobertura territorial del CENSO y la información socioeconómica de los hogares de la encuesta. La Figura 2 muestra el modelo relacional "CENSO-CASEN", que representa los datos centralizados de CENSO en "1" con los datos "2" pertenecientes a CASEN.

Figura 2: Modelado de datos relacional entre CENSO 2002 y la encuesta CASEN 2003



Fuente: elaboración propia.

La estructura del modelo de datos “CENSO-CASEN”, alojada en un almacén de datos intermedio; se pobló mediante un proceso ETL con la información de cada instrumento (datos en archivos SPSS), desde donde se extrajeron los datos para luego alojarlos en un almacén intermedio. El sub-proceso de transformación concilió los códigos de localización geográfica en ambos instrumentos dando paso a la implementación del matching espacial; proceso que dio paso a consultas combinadas sobre el almacén de datos intermedio, cuya estructura se genera gracias a la existencia de variables homogéneas a ambos instrumentos que permite parrear observaciones que comparten características similares; es decir, encontrar personas del CENSO con características similares a otras de la encuesta, para que se pueda imputar luego; la ubicación espacial de los primeros a los segundos dentro de la comuna. La forma de hacer esta asociación se realiza en dos etapas. La primera etapa consiste en la aplicación de un pareo exacto, que busca, para cada observación perteneciente a la encuesta con características definidas por un vector de variables comunes, una observación perteneciente al CENSO que posea las

mismas características; transfiriendo así la ubicación espacial del ingreso de los hogares disponibles en la encuesta a nivel de comunas pero en un lugar específico dentro de esa comuna. La segunda etapa, viene dada por la obtención del valor de la variable (ingreso de los hogares) en pequeñas áreas, que se obtiene como el promedio del ingreso de las observaciones de la encuesta y que comparten la misma información en CENSO. Esta etapa, permite pasar de un “ÚNICO” ingreso medio representativo de los hogares de esa comuna, a un valor del ingreso que cambia de acuerdo a las condiciones económicas de cada distrito dentro de la comuna.

Por tanto, con el matching se logra la desagregación de los datos de encuesta en el micro territorio con una distribución homogénea de los hogares en el territorio, a una distribución urbana segregada en algunos casos; de acuerdo a su condición socioeconómica.

SISTEMAS DE INFORMACIÓN GEOGRÁFICA

Base de datos espacial

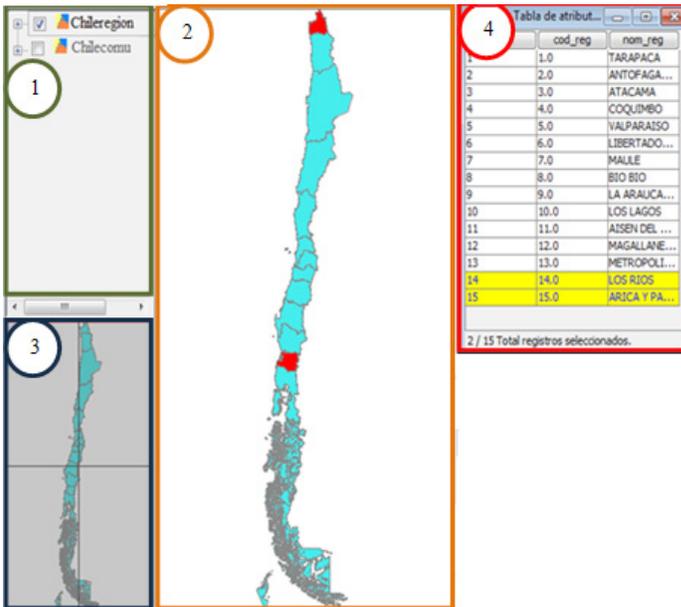
Para la observación geográfica de la información generada entre CENSO y CASEN, se incorporó un sistema administrador de bases de datos espaciales PostgreSQL con el módulo de PostGIS, el cual añadió soporte de objetos geográficos a la base de datos relacional; a través de campos geométricos en las tablas de datos, tablas con proyecciones espaciales y funciones entre otras, permitiendo de este modo generar bases de datos espaciales. Con la extensión para el manejo y análisis de objetos SIG, se generó una base de datos espacial, donde se crearon esquemas tanto para el almacenamiento de los datos depurados en la etapa de pre-procesamiento y para el alojamiento de la cartografía georreferenciada, como para el hospedaje de las tablas que contienen los indicadores, los que otorgan información complementaria a la cartografía a distintos niveles territoriales.

SIG Desktop

Generada y estructurada la base de datos espacial, se realizó mediante gvSIG, la carga de la cartografía del territorio chileno. Para esto, se utilizó como fuente de datos, la información alojada en archivos en formato Shapefile que contienen la localización de elementos geográficos y sus atributos asociados en un formato de representación vectorial (Universidad de Alcalá, 2012). Con gvSIG se generaron mapas específicos de regiones, provincias y distritos, además del mapa

de Chile (quince regiones), previa carga en la base de datos espacial y asignación del sistema de proyección para representar las coordenadas sobre un plano. La Figura 3 muestra la cartografía regional de Chile (división regional de 2007) que fue generada a partir de la cartografía comunal de Chile según división regional de 1976. La imagen está seccionada en cuatro: la sección “1” muestra las capas “chilecomu” y “chileregion”, que representan la cartografía fraccionada en comunas y regiones respectivamente. La sección “2” contiene la cartografía de la capa seleccionada. La sección “3”, el posicionamiento en el territorio y la sección “4”, la tabla de atributos de regiones, remarcándose en amarillo las regiones de “Arica y Parinacota” y “Los Ríos”, cuya relación gráfica queda expresada con color rojo en el mapa de Chile sección “2”.

Figura3: Cartografía regional de Chile

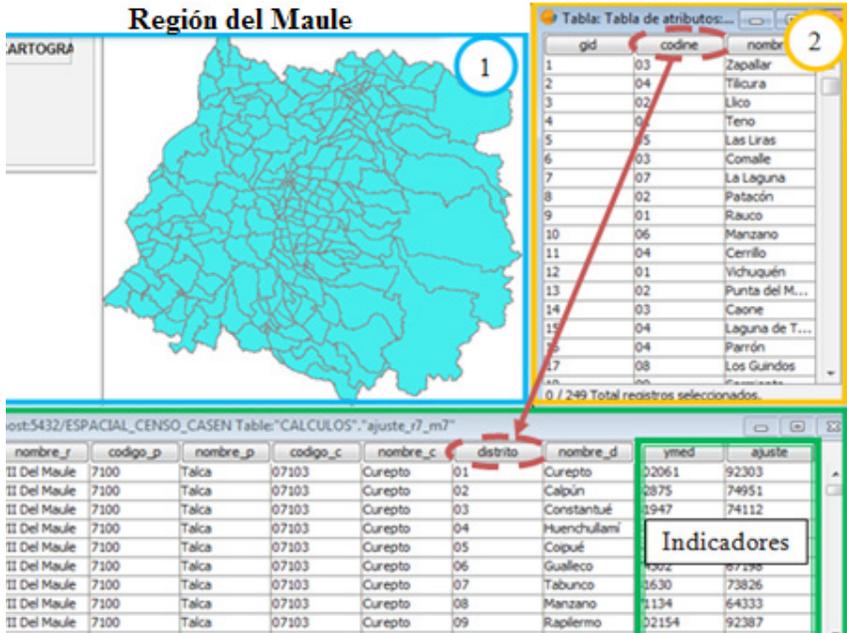


Fuente: elaboración propia.

Generada y cargada la cartografía en la base de datos espacial, se llevó a cabo la elaboración de indicadores para el territorio nacional, a partir de los instrumentos de medición CENSO y CASEN, como también del matching generado a partir de éstos. La visualización de los indicadores en gvSIG, se obtuvo mediante tablas de atributos relacionados con las tablas que contenían la cartografía, través de las claves utilizadas en el territorio (ej.: código de comunas o distritos, entre otros). Por ejemplo: la Figura 4 de gvSIG que muestra la Región del Maule en

cuatro; muestra en la sección 1 un mapa de los distritos y en la sección 2 su respectiva tabla de atributos; en la sección 3 la tabla de atributos con los indicadores respectivos y que mediante las funcionalidades de gvSIG, se puede representar visualmente a través del código equivalente entre éstas (línea entre sección 2 y 3).

Figura 4: Incorporación de indicadores en gvSIG

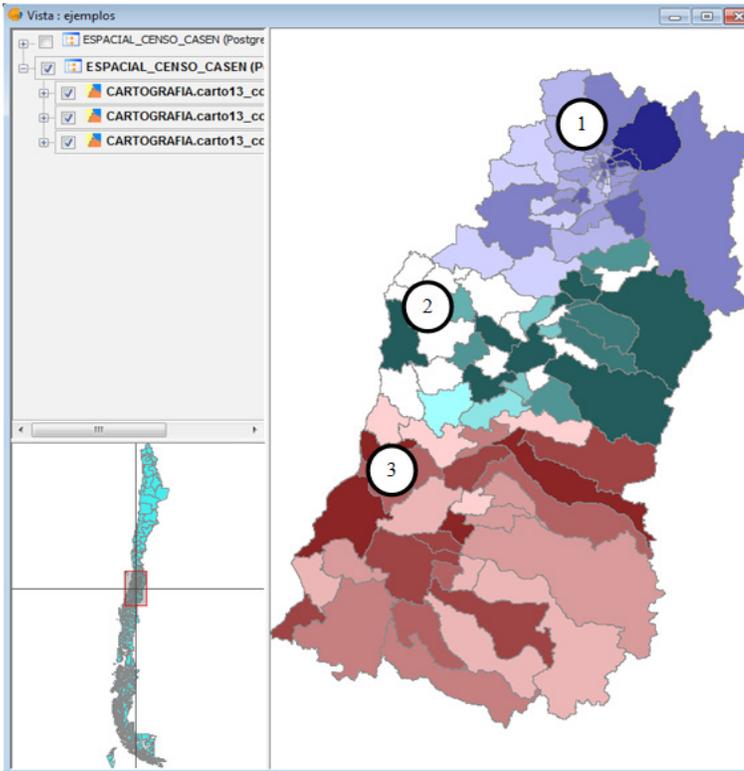


Fuente: elaboración propia.

La herramienta gvSIG permitió, entre otras cosas, analizar de forma gráfica los indicadores generados y asociados con el territorio, mediante el uso de propiedades que este aplica a las capas. Por ejemplo, la Figura 5 graficada con gvSIG, muestra los intervalos naturales (clasificación por rango de valores) para los ingresos medios por comuna de las regiones VI, VII y XIII, representadas por los valores “1”, “2” y “3” respectivamente. Para la construcción de esta, se extrajeron valores directamente de la encuesta CASEN y se presentaron al mayor nivel de granularidad entregado por esta encuesta. En el mapa, las regiones se analizan de forma independiente, donde los colores más densos representan los ingresos medios más elevados, mientras que los terrenos de color blanco no representan valor, debido a que no fueron encuestados. Esta representación gráfica permite analizar la distribución de ingresos en el territorio y distinguir los diversos estratos sociales

existentes en las regiones.

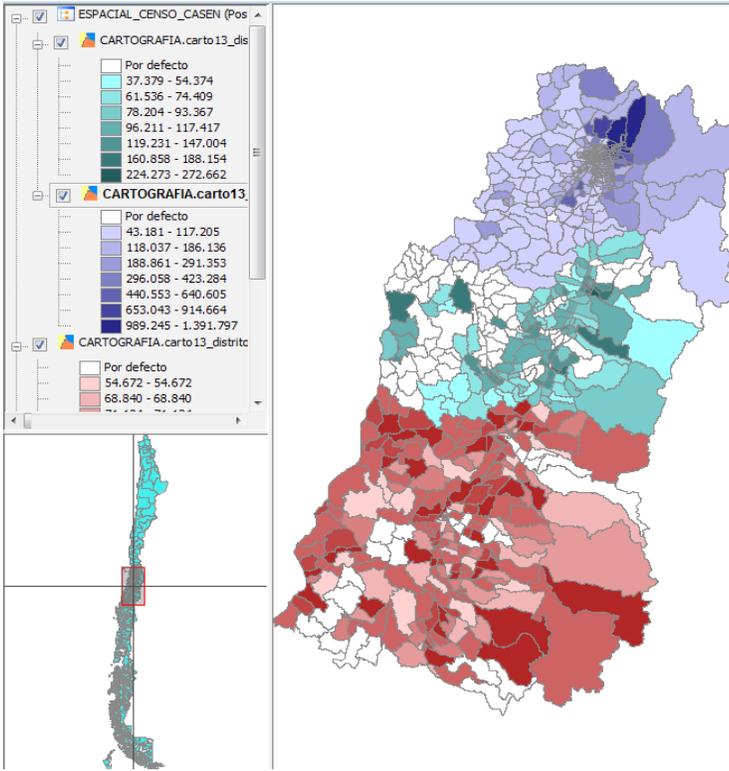
Figura 5: Ingresos medios por comuna para las regiones XIII, VI y VII



Fuente: Elaboración propia.

En la Figura 6 se pueden apreciar las mismas regiones pero en niveles micro territoriales (nivel de distritos) obtenidos a través del matching entre los instrumentos de medición CENSO y CASEN permitiendo con ello, observar en mayor detalle la distribución de los ingresos de la región. Por otro lado; con el matching se tiene acceso a un nivel de distritos lo cual permite aumentar el número de observaciones facilitando con ello la estimación de datos ausentes mediante el interpolador kriging; técnica proveniente de la geoestadística que permite estimar el valor ausente de la variable en el territorio.

Figura 6: Ingresos medios por distrito para las regiones metropolitanas, VI y VII



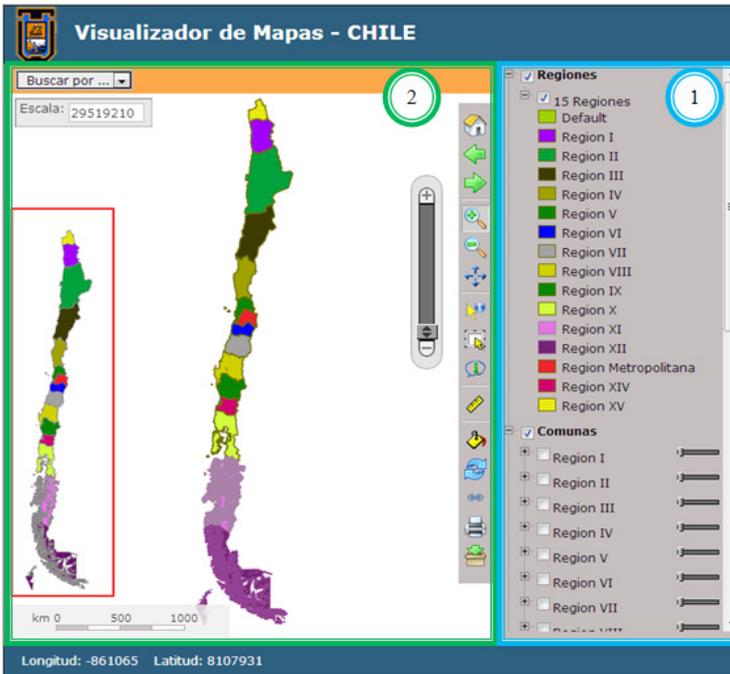
Fuente: Elaboración propia.

Plataforma de visualización cartográfica

Finalmente, para ampliar la capacidad de análisis de los datos, se generó una aplicación Web para la publicación de datos espaciales, mediante la tecnología **Internet Map Server** tanto en modo vectorial como ráster (mapa de bits). Lo anterior, se implementó en una plataforma constituida por un servidor Web, un servidor de mapas, y un conjunto de librerías para el uso y tratamiento de información geográfica. Para la carga de la cartografía en el servidor de mapas, se prepararon los datos a utilizar mediante "Mapfile", archivo principal de configuración de MapServer cuyo propósito fundamental es definir: las capas del mapa a generar, el estilo con que se mostrarán, la simbología, el sistema de referencia y el formato de las imágenes, entre otras (Mapserver, 2012). En este archivo se definieron las consultas a la base de datos espacial para la importación de la información geográfica. Parte del código

utilizado para la elaboración de la cartografía en el archivo Mapfile se realizó en gvSIG, mediante de la incorporación de un plugin para publicaciones Web. Luego, utilizando las funcionalidades y múltiples configuraciones proporcionadas por el Framework p.mapper, se generó la plataforma Web mediante la cual se visualizó y manejó la cartografía importada y configurada en el archivo Mapfile. La Figura 7 presenta la plataforma de visualización de mapas desarrollada a través de los componentes previamente mencionados. La plataforma se divide en dos secciones: **Capas de información geográfica**: ubicada a la derecha de la figura y destacada con el número “1”, que contiene la cartografía (divisiones territoriales) utilizada seccionada en capas de información con sus símbolos correspondientes y agrupada según su tipo (comunas, distritos e ingresos medios de la región, entre otros) y **Ventana de visualización**: ubicada a la izquierda de la pantalla y destacada con el número “2”. En esta sección se muestran los elementos cartográficos respectivos a las capas seleccionadas de información geográfica.

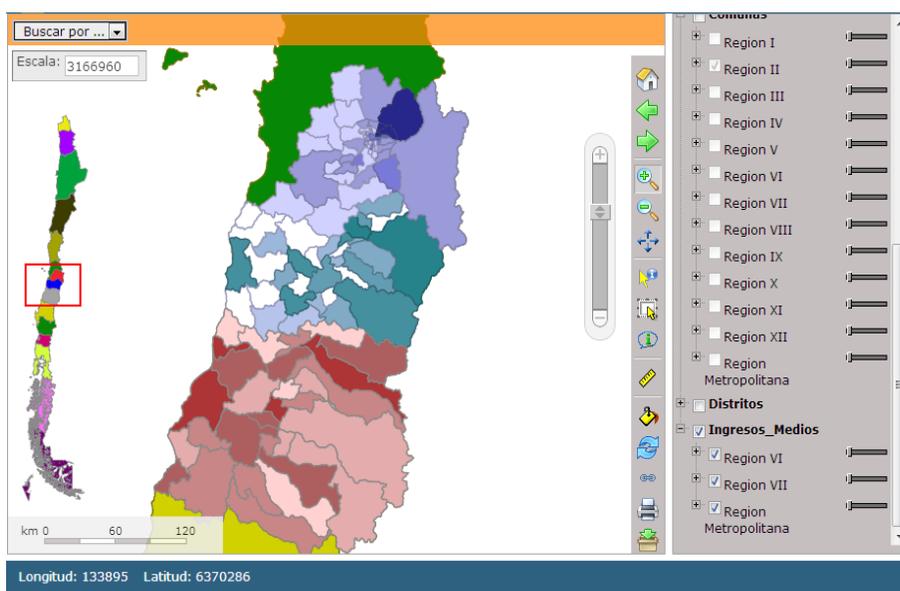
Figura 7: Plataforma de visualización de mapas



Fuente: Elaboración propia.

Algunos componentes y funcionalidades de la plataforma de visualización de mapas son la **Escala** (tamaño de visualización de la cartografía), **Barra de herramientas** (manejo de la escala de visualización del mapa, medición de distancias, generación de links de acceso a ciertas ubicaciones y descarga de las vistas entre otros) y **Mapa de referencia** (visualización de la ubicación en la cual se encuentra posicionado el usuario, destacando ésta con un rectángulo de color rojo, ver Figura 7, sección 2). La plataforma de visualización de mapas, además incorpora capas de indicadores, como por ejemplo: indicadores de ingresos medios para las comunas y distritos tratados en la sección SIG Desktop (ver Figura 8).

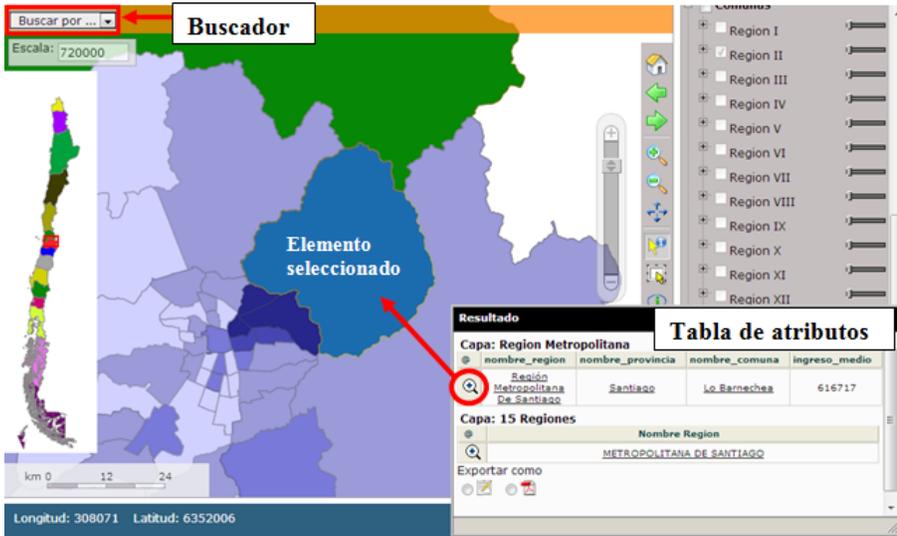
Figura 8: Visualización de indicadores



Fuente: Elaboración propia.

La Figura 9 muestra la tabla de atributos que contiene información específica del conjunto seleccionado, la cual permite visualizar más en detalle la ubicación del territorio perteneciente al conjunto de elementos. Además, la plataforma cuenta con un buscador, en el cual se debe seleccionar la capa que se desea examinar. La plataforma también da la opción de imprimir informes y descargar, en diversas escalas, la información respectiva a la capa seleccionada en formato PDF o CSV. Además, permite al usuario, generar enlaces con una posición específica en el mapa para ser visualizados por otro usuario.

Figura 9: Buscador y tabla de atributos



Fuente: Elaboración propia

CONCLUSIONES

En este trabajo hemos pretendido mostrar que la utilización de los SIG en la observación espacial de eventos socioeconómicos, puede no solo contribuir con la representación geográfica de variables como el ingreso medio de los hogares, sino que mediante la integración de datos de CENSO y encuesta se facilita la posibilidad de mejorar la precisión en las estimaciones de datos inexistentes sobre la condición socioeconómica de sus habitantes en el territorio, así como la observación de zonas de rezago económico a fin de generar, por ejemplo, programas de asistencia pública. Nuestra propuesta pretende llamar la atención a los investigadores sobre la utilidad de los SIG en la observación de los eventos socioeconómicos. Mediante un proceso de ingeniería inversa, se construyó un modelo relacional de CENSO, al que se aplicó un matching espacial con los datos de la Encuesta a fin de extrapolar la información socioeconómica de los hogares a unidades micro territoriales no disponibles en ella (distritos o divisiones intra comunales). La utilización de la herramienta SIG Desktop permitió entre otras cosas, llevar a cabo, tanto la carga de la cartografía del territorio chileno en la base de datos espacial, como la generación, edición y análisis de los datos, entregando un repositorio de cartografía digital del territorio nacional, y sus indicadores asociados relacionados con el CENSO y encuesta. Por último, la implementación de la plataforma de visualización de

mapas se realizó mediante la tecnología Internet Map Server, con el fin de permitir a los usuarios visualizar, consultar y analizar información geográfica a través de la red, tanto en modo vectorial como ráster. Para ello, se levantó una plataforma constituida por el servidor Web, el servidor de mapas, y un conjunto de librerías para el uso y tratamiento de información geográfica. Esta plataforma, en comparación con una herramienta “SIG Desktop”, es más acotada, lo cual permite brindar sólo información de interés a un grupo específico de usuarios. A modo general, esta solución permitirá a los expertos obtener una mejor comprensión de los datos espaciales y resolver problemas específicos relacionados con el territorio de forma más eficiente, debido a que esta brinda una superior perspectiva en el proceso de toma de decisiones.

BIBLIOGRAFÍA

- Simitsis, P. Vassiliadis and T. Sellis (2005). “Logical Optimization of ETL Workflows”. *IEEE Transactions on Knowledge and Data Engineering*. 17 (10), 1404–1419.
- C. Agostini, P. Brown y D. Góngora (2008). Distribución Espacial de la Pobreza en Chile. *Estudios de Economía*. 1(35), 79-110.
- C. Agostini, P. Brown y A. Román, (2010). “Estimando indigencia y pobreza indígena regional con datos censales y encuestas de hogares”. *Cuadernos de Economía* (47), pp. 125-150.
- C. Elbers, J.O. Lanjouw y P. Lanjouw (2003). “Micro-Level Estimation of Poverty and Inequality”. *Econometrica*, 71 (1), 355-364.
- C. Elbers y P. Lanjouw (2012). “Captando dinámicas de pobreza en el nivel sub-nacional: Seguimiento de cambios por medio de estimaciones para áreas pequeñas” en *Los dilemas territoriales del desarrollo en América Latina*. (Modrego, F. y J.A. Berdegué, eds.). Buenos Aires: Editorial Teseo.
- D. Paredes y P. Aroca (2008). “Metodología para estimar un índice regional de costo de vivienda en Chile”. *Cuadernos de Economía* (45), pp. 129-143.
- F. Modrego, E. Ramírez y A. Tartakowski (2008). “La heterogeneidad espacial del desarrollo económico en Chile: radiografía a los cambios en bienestar durante la década de los noventa por estimaciones en pequeñas áreas”. Documento de trabajo N° 9 Programa Dinámicas Territoriales Rurales Rimisp –

Centro Latinoamericano para el Desarrollo Rural.

- F. Modrego, E. Ramírez, A. Tartakowski y E. Jara (2012). "La heterogeneidad territorial del desarrollo en la década de oro de la economía chilena" Los dilemas territoriales del desarrollo en América Latina. (Modrego, F. y J.A. Berdegú, eds.). Buenos Aires: Editorial Teseo.
- GEOINFO (s.f.) "Sistemas de Información Geográficos" Disponible en <http://www.geoinfo.cl/pdf/sig.pdf>
- Instituto Nacional de Estadísticas (2002). Informe: CENSO 2002 RESULTADOS. Población País - Región. - Volumen I. Chile.
- J. ConesaCaralt y J. Curto Díaz (2010) Introducción al Business Intelligence. Barcelona: Editorial UOC.
- Ministerio de Planificación (MIDEPLAN) (2005). Informe "marco metodológico CASEN2003". Chile.
- MapServer (s.f.) "Open source web mapping". Disponible en: <http://mapserver.org/es/about.html#about> [consultado: 14/09/2012].
- N. Minot y B. Baulch (2005). "Spatial patterns of poverty in Vietnam and their implications for policy". *Food Policy*, 30 (5-6), 461-475.
- N. Sillero P. Tarrasopublicacion (2010) Tesis: Free GIS for herpetologists: free data sources on Internet and comparison analysis of proprietary and free/open source software. Centro de Investigación en Ciencias Geo-Espaciales (CICGE) da Universidad de Porto. R. Campo Alegre. Publicación Magister. pp. 64-81.
- OSGeo Project (s.f.) "What is PostGIS?" Disponible en: <http://postgis.refractor.net/> [consultado: 11/10/2012]
- P. Leija Luna (2010). Tesis: "Sistema de Información Geográfica para la ayuda de toma de decisiones en políticas sociales" Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional Departamento de Computación. Maestra Ciencias en la Computación, pp. 9-11.
- PostgreSQL (s.f.) "Sobre PostgreSQL" Disponible en: <http://www.postgresql.org/>

org.es/sobre_postgresql [consultado: 11/10/2012]

- R. Kimball, L. Reeves, M. Ross and W. Thornthwaite (1998). The Data Warehouse Lifecycle Toolkit: Expert Methods for Designing, Developing, and Deploying Data Warehouses. John Wiley and Sons.
- Universidad de Alcalá/Departamento de Geografía y Geología (s.f.) "Practica Vectorial formato Shapefile" Disponible en: http://www.geogra.uah.es/gisweb/practica-vectorial/Formato_Shapefile.htm [consultado: 04/09/2012].
- W. H. Inmon (2002). Building the Data Warehouse. Wiley. 3a. ed.