Construcción de un mercado de datos para el almacenamiento de lesiones de causa externa*[1]

[Building a Data Mart for the storage of injuries from external causes]

Ricardo TIMARÁN PEREIRA[2], Alexander BARÓN[3], Gonzalo HERNÁNDEZ[4]

Recibo: 20.11.2013 – Aprobación: 15.03.2014

Resumen

En este artículo se presenta el resultado de la segunda fase del proyecto de investigación que tiene como objetivo detectar patrones delictivos con técnicas de minería de datos en el Observatorio del Delito del municipio de Pasto (Colombia). En esta fase se construyó un mercado de datos que almacenará los datos históricos, limpios y transformados sobre las lesiones de causa externa fatal y no fatales registradas en este Observatorio en un periodo de 10 años. Este mercado permitirá realizar análisis multidimensional y soportar la detección de patrones delictivos que facilite a los organismos gubernamentales y de seguridad tomar decisiones eficaces en lo relacionado a la implementación de planes de prevención de delitos y seguridad ciudadana.

* **Modelo para la citación de este artículo**:
TIMARÁN PEREIRA, Ricardo; BARÓN, Alexander & HERNÁNDEZ, Gonzalo (2014). Construcción de un mercado de datos para el almacenamiento de lesiones de causa externa. En: Ventana Informática No. 30 (ene-jun). Manizales (Colombia): Facultad de Ciencias e Ingeniería, Universidad de Manizales. p. 67-79. ISSN: 0123-9678

1 Artículo de investigación científica y tecnológica proveniente del proyecto *Detección de Patrones Delictivos con Técnicas de Minería de Datos en el Observatorio del Delito del Municipio de Pasto*, ejecutado en el periodo *10/2011-09/2013*, financiado por COLCIENCIAS e inscrito en el *Grupo de Investigación Aplicada en Sistemas- GRIAS* de la *Universidad de Nariño.*

2 PhD. en Ingeniería, MSc. en Ingeniería, Esp. en Multimedia e Ingeniero de Sistemas y Computación. Director grupo de investigación GRIAS, Profesor Asociado, Departamento de Sistemas, Facultad de Ingeniería, Universidad de Nariño, Pasto (Colombia). Correo electrónico: ritimar@udenar.edu.co

3 MSc. en Ingeniería Informática, Esp. en Desarrollo de Software. Profesor Asociado, Departamento de Sistemas, Facultad de Ingeniería, Universidad de Nariño, Pasto (Colombia). Correo electrónico: abaron_98@udenar.edu.co

4 MSc. en Ingeniería de Sistemas y Computación, Esp. en Redes y Servicios Telemáticos. Profesor Asociado, Departamento de Sistemas, Facultad de Ingeniería, Universidad de Nariño, Pasto (Colombia). Correo electrónico: GonzaloHernandez@udenar.edu.co

Palabras Clave: *Mercado de Datos, Patrones Delictivos, Análisis Multidimensional, Minería de Datos*

Abstract

This paper presents the results of the second phase of the research project that aims to detect crime patterns with data mining techniques in the Crime Observatory of the municipality of Pasto (Colombia). At this stage, it has been built a data mart that will store historical, clear and processed data about injuries from fatal and non-fatal external causes recorded at the Observatory over a period of 10 years. This data mart will support multidimensional analysis in order to detect crime patterns which aim to facilitate the effective decision making in relation to the implementation of crime prevention plans and public safety by the government and security agencies.

Keywords: *Data Mart, Criminal Patterns, Multidimensional Analysis, Data Mining.*

Introducción

La violencia se considera un problema de salud pública (OPS, 2002,4). Este hecho es aceptado por la Organización Panamericana de la Salud, OPS, desde 1993 y por la Organización Mundial de la Salud, OMS, desde 1996. La OMS define la violencia como «*el uso deliberado de la fuerza física o el poder, ya sea en grado de amenaza o efectivo, contra uno mismo, otra persona o un grupo o comunidad, que cause o tenga muchas probabilidades de causar lesiones, muerte, daños psicológicos, trastornos del desarrollo o privaciones*» (OPS, 2002, 5).

Según el informe Mundial sobre la Violencia y la Salud (OPS, 2002), más de 1,6 millones de personas en todo el mundo pierden la vida violentamente, cada año. Una de las principales causas de muerte en la población de edad comprendida entre los 15 y los 44 años y la responsable del 14% de las defunciones en la población masculina y del 7% en la femenina es la violencia (OPS, 2002, 1). Este fenómeno, cada año impone a las economías de los países del mundo, una gran carga financiera, de miles de millones de dólares, por concepto de atención sanitaria, gastos judiciales, policiales y pérdida de productividad.

En el estudio sobre Observatorios de Violencia (OPS, 2008a), realizado por el Instituto de Investigaciones y Desarrollo en Prevención de Violencia y Promoción de la Convivencia Social ,CISALVA de la Universidad del Valle (Colombia), en países de América Latina tales como Colombia,

Venezuela, El Salvador, Guatemala y Honduras, en estimativos de los años 2003 y 2005, se reportaron altos índices de homicidio, con tasas iguales o superiores a 29 homicidios por 100.000 habitantes (OPS, 2008a, 7).

En este estudio se reconoce que:

> «*A pesar del problema documentado, las cifras disponibles no son de buena calidad. La mayor parte de las veces es necesario desarrollar estudios específicos para conocer el problema o son datos oficiales que presentan subregistros o no coinciden al compararlos con otras fuentes de información. Por lo tanto, es necesario recopilar datos e información que permitan mejorar el conocimiento sobre la magnitud y las características de los hechos, orientar estudios para identificar los factores que inciden en la presencia o no del evento, evaluar políticas e intervenciones y hacer difusión de las mismas*» (OPS, 2008a, 7).

La ausencia de información confiable y oportuna es un problema en América Latina. Este hecho restringe la posibilidad de avanzar en la identificación de la magnitud y características de las diferentes formas en que se expresa la violencia, así como el seguimiento y evaluación de los programas y proyectos para su prevención y control (OPS, 2008b, 5).

Una de las estrategias implementadas en vigilancia en salud pública, corresponde a los Observatorios de Muertes de Causa Externa (OPS, 2008b, 5), los cuales se han instaurado para el seguimiento y análisis en el nivel local (municipal) en casos de mortalidad por causa externa como: homicidios, suicidios, eventos de tránsito y muertes no intencionales. En Colombia, según OPS (2008b, 5) se cuenta con diversas experiencias en los entornos municipal y departamental.

En el municipio de Pasto, el observatorio de muertes por causa externa, denominado Observatorio del Delito, nace en el segundo semestre del año 2002, como resultado de un proyecto conjunto con el Programa Colombia de la Universidad de Georgetown. El Observatorio del Delito, es formalizado mediante el Acuerdo 022 de noviembre de 2002, del Concejo Municipal de Pasto, como lo aseguran OPS (2008b), y Betancourt-Salazar (2005, 3).

Actualmente, se desarrolla el proyecto de investigación que tiene como objetivo descubrir patrones delictivos con técnicas de minería de datos a partir de los datos almacenados en el Observatorio del Delito del municipio de Pasto (Colombia). Este proyecto consta de tres fases. Como resultado de la primera fase, el Observatorio del Delito consolida su sistema de vigilancia de eventos violentos con la implementación

de SIGEODEP, un sistema de información georreferenciado que le permite disponer de información confiable, oportuna, de buena calidad y representativa de las lesiones de causa externa que ocurren en el municipio de Pasto, para que sus autoridades tomen decisiones acertadas en materia de prevención, atención y control de la violencia y la accidentalidad (Timarán et al., 2012, 2).

Como resultado de la segunda fase de este proyecto, se integró a SIGEODEP, un mercado de datos que almacena los datos históricos, limpios y transformados sobre las lesiones de causa externa fatal y no fatales registradas en el Observatorio del Delito en un periodo de 10 años. Con el fin de extraer, transformar y cargar los datos limpios al mercado de datos, se desarrolló un sistema ETL (*Extraction, Transformation and Loading*) y se acopló a SIGEODEP. Este mercado permitirá realizar procesamiento analítico en línea con herramientas OLAP (*Online Analytical Processing*) y soportar la detección de patrones delictivos con herramientas de minería de datos (Elmasri & Navathe, 2007,851). Estos resultados facilitarán a los organismos gubernamentales y de seguridad tomar decisiones eficaces en lo relacionado a la implementación de planes de prevención de delitos y seguridad ciudadana.

El resto del artículo se organiza de la siguiente manera. En la sección 1, se presenta los conceptos básicos sobre bodegaje de datos. En la sección 2, se describe la metodología utilizada para la construcción del mercado de datos. En la sección 3, se presentan los resultados del proceso de construcción del mercado de datos y finalmente, en la última sección se presenta las conclusiones y trabajos futuros.

1. Fundamento teórico

1.1 Almacén de datos

La definición universalmente aceptada para un almacén de datos (*datawarehouse*) es la propuesta por Willian Inmon en 1980, como una colección de datos orientada a temas, integrada, variable en el tiempo y no volátil para soportar la toma de decisiones estratégicas, coinciden Imhof, Galemmo & Geiger (2003, 9), Inmon (2005, 31) y Connolly & Begg (2005, 1039). Por otra parte, Hernández, Ramírez & Ferri (2005, 47) y Rainardi (2008, 1), definen un almacén de datos como un conjunto de datos históricos, internos o externos y descriptivos de un contexto o área de estudio, que están integrados y organizados de tal forma que permiten aplicar eficientemente herramientas para resumir, describir y analizar los datos con el fin de ayudar en la toma de decisiones estratégicas. Según Ponniah (2010, 13), la ventaja

fundamental de un almacén de datos es su diseño específico y su separación de la base de datos transaccional. Este hecho facilita el análisis de los datos en tiempo real (OLAP) y no disturba el OLTP de las bases de datos originales.

1.2 Modelo multidimensional

Es el modelo conceptual de datos propuesto por Kimball (Kimball & Ross, 2002, 16) para los almacenes de datos. En el modelo multidimensional una actividad que es objeto de análisis se organizan en torno a los hechos, que tienen unos atributos o medidas que pueden verse en mayor o menor detalle según ciertas dimensiones. Cada dimensión tiene una estructura jerárquica pero no necesariamente lineal. La forma como se organizan los hechos y sus dimensiones determina el esquema multidimensional del almacén de datos. En un esquema en estrella las dimensiones tienen una estructura jerárquica lineal y en un esquema de estrella jerárquica o copo de nieve existen caminos alternativos en las dimensiones.

1.3 Mercado de datos

No es posible recopilar toda la información de una organización en un almacén de datos con un único esquema en estrella o copo de nieve. Por el contrario, es necesario que para cada departamento o sección de la organización se construya con un esquema específico. El conjunto de esquemas forman el almacén de datos y a cada esquema se le denomina mercado de datos o *datamart* (Kimball & Ross, (2002), Connolly & Begg, (2005)).

1.4 Implementación de un almacén de datos

Un almacén de datos puede implementarse utilizando dos tipos de esquemas físicos: ROLAP o MOLAP, señalan Hernández, Ramírez & Ferri (2005, 55). En un sistema ROLAP, el almacén de datos se implementa sobre tecnología relacional, pero disponen de algunas facilidades para mejorar el rendimiento (índices de mapas de bits, índices de JOIN). En un sistema MOLAP, el almacén de datos se implementa sobre estructuras basadas en matrices multidimensionales y técnicas de compactación que favorecen el rendimiento del almacén.

2. Metodología

Siguiendo la metodología de desarrollo de un almacén de datos, presentada por Anaya (1996, 96), Inmon (2000, 17), Kimball & Ross (2002, 30) y Rainardi (2008, 50), que se muestra en la figura 1, se construyó el mercado de datos para soportar la detección de patrones delictivos.

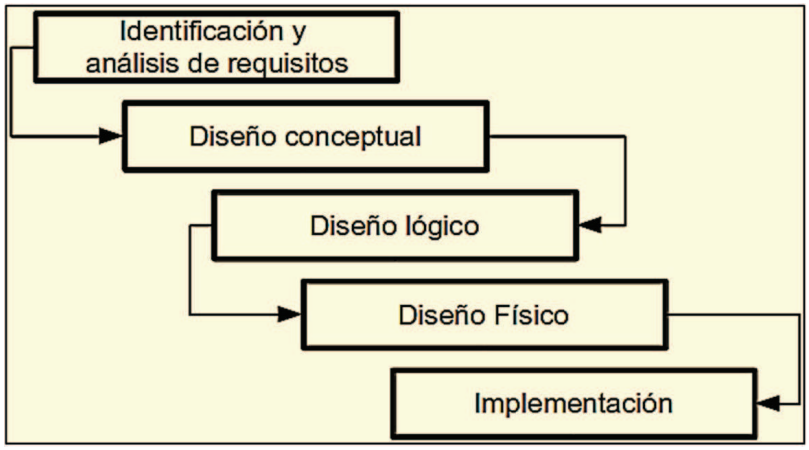


Figura 1. Metodología de diseño

2.1 Identificación y análisis de requisitos

Una interacción directa con los usuarios del futuro sistema siempre será un factor crítico en el éxito de una implementación de bodegas de datos. Afortunadamente, mucho del trabajo de identificación y análisis de requisitos ya se había abordado durante la primera fase del proyecto. Los canales de comunicación e interacción con los funcionarios del Observatorio del Delito y los miembros del grupo de desarrollo se mantuvieron de manera similar durante el desarrollo de SIGEODEP y el diseño del almacén de datos.

Los requisitos de información y los diferentes informes solicitados por el Observatorio se conocían de antemano. A través de reuniones mensuales se compartió con los funcionarios del Observatorio conceptos básicos sobre el modelado de almacenes de datos y varias versiones preliminares del diseño del almacén con el objetivo de verificar que todas las necesidades de información estuvieran cubiertas.

2.2 Diseño conceptual

Se construyeron diferentes alternativas de diseño usando diagramas Entidad – Relación que se discutieron con los usuarios. La figura 2 muestra el diseño del prototipo de un *datamart* para el análisis de la información general del Observatorio sobre el tipo de evento, el lugar y momento de una lesión.

2.3 Diseño lógico

Una vez se validó el diseño conceptual, se procedió a formular el diseño lógico del almacén de datos. En esta instancia, se tomaron las diferentes tablas de hechos y sus dimensiones asociadas y se diagramó sus

respectivas jerarquías y niveles bajo un modelado multidimensional. Por ejemplo, la figura 2 ilustra el diseño lógico para el *datamart*.

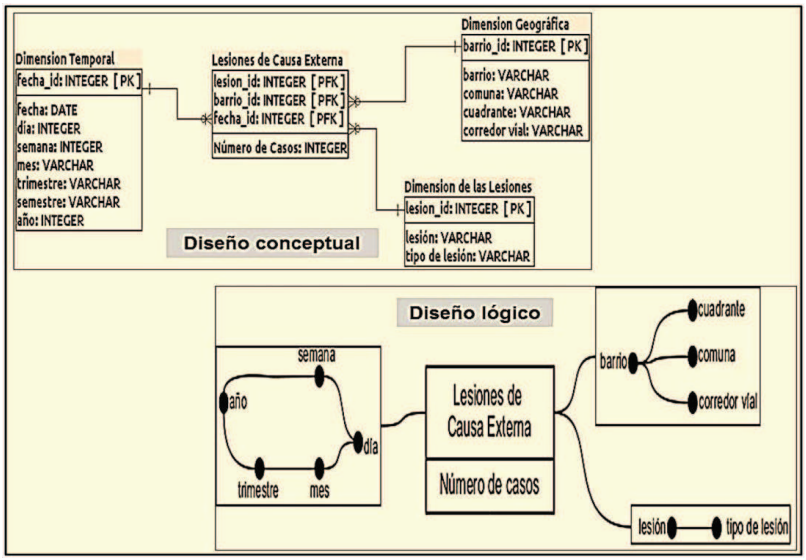


Figura 2. Diseños conceptual y lógico del prototipo de datamart.

2.4 Diseño físico

Posterior al diseño lógico, se continuó con el diseño físico del almacén de datos. Aquí se definió bajo que arquitectura y esquema se implementarán los *datamarts* del almacén. Se decidió seguir un esquema en estrella e implementarlo bajo una arquitectura ROLAP acoplada a *PostgreSQL*.

Para el diseño de los procesos ETL se utilizó el componente de integración de datos (PDI – *Pentaho Data Integration*) de la suite de inteligencia de negocios *Pentaho*. La figura 3 muestra una de las rutinas diseñadas para cargar los registros desde el sistema transaccional hacia el prototipo del *datamart*.

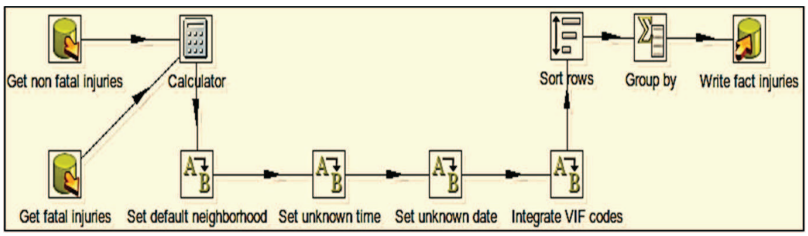


Figura 3. Diseño de un proceso ETL para alimentar el prototipo de datamart.

2.5 Implementación

Durante la implementación del mercado de datos se hizo uso de diversas herramientas libres como SQL *Power Architect* y *PostgreSQL.* Para la construcción de los procesos ETL, construcción de cubos e implementación de prototipos se utilizó el conjunto de herramientas proveído por la suite de inteligencia de negocios *Pentaho* en su edición comunitaria. Como se mencionó anteriormente, *Pentaho Data Integration* fue usado para la creación de transformaciones, extracción y carga de datos desde el sistema transaccional al mercado de datos. Se utilizó *Pentaho Schema Workbench* (PSW) para la implementación de cubos OLAP y *Pentaho Analysis Services* (PAS) para la visualización de dichos cubos.

3. Resultados

Actualmente se ha implementado un prototipo del almacén de datos del Observatorio del Delito y una serie de *datamarts* para cada uno de los eventos que vigila el Observatorio. Igualmente se cuenta con cubos OLAP y visualizaciones que facilitan el cruce de variables entre las tablas de hechos definidas en el diseño y sus dimensiones. A continuación, se describe en más detalle la construcción de las diferentes tablas de hechos, dimensiones compatibles y particulares, métricas y visualizaciones implementadas.

3.1 Tablas de hechos

El observatorio del Delito monitorea tres líneas de vigilancia: lesiones de causa externa fatales, lesiones de causa externa no fatales y violencia intrafamiliar. Dentro de las lesiones de causa externa fatales se pueden identificar cuatro eventos: homicidios, suicidios, muertes no intencionales y muertes en accidentes de tránsito. Cada uno de estos cuatro eventos se implementó como una tabla de hechos dentro del mercado de datos.

De manera similar, la línea de vigilancia de lesiones de causa externa no fatales registra datos sobre cuatro tipos de lesiones: violencia interpersonal, lesión en accidente de tránsito, lesiones autoinflingidas y lesiones no intencionales. Al igual que en el caso de las lesiones fatales, cada evento se constituyó como una tabla de hechos.

Finalmente, la línea de vigilancia de violencia intrafamiliar se centra en ese tipo de eventos en particular y se implementó como una tabla de hechos independiente. A la par con las tablas de hechos definidas para cada evento, se decidió implementar tres tablas adicionales: dos que resuman los datos comunes tanto para la línea de vigilancia de lesiones de causa externa fatal como no fatal y una tabla de hechos general que recopile las principales características que son comunes en todos los eventos, por ejemplo: el lugar, fecha y características de la víctima.

3.2 Dimensiones compatibles

Como en la mayoría de las implementaciones de almacenes de datos, las dimensiones temporal y geográfica son comunes para todos los eventos. La dimensión temporal se dividió a nivel de fecha y hora con diferentes niveles de agregación. La dimensión de la fecha posee jerarquías a nivel de día, semana, mes, trimestre, semestre y año. Por su parte, la dimensión de la hora se concentra en cuando y en qué parte del día ocurrió la lesión (madrugada, mañana, tarde o noche).

La dimensión geográfica también tiene diferentes niveles de agregación. Cada evento registrado se asocia al barrio en que ocurrió la lesión. A partir del barrio, es posible calcular agregaciones a nivel de comuna, cuadrante y corredor vial. Vale la pena aclarar que en el sector rural se utiliza la convención vereda/corregimiento de manera análoga a como se utiliza barrio/comuna en el contexto urbano.

Junto con las dimensiones temporal y geográfica, otra dimensión compatible a todas las lesiones es la dimensión de la víctima. En esta dimensión se recopilaron las características más relevantes de cada individuo afectado por una lesión. Entre ellas se cuenta: edad de la víctima en diferentes niveles (grupos quinquenales, grupos de edades y si es mayor de edad o no), género de la víctima, ocupación y barrio de residencia.

3.3 Dimensiones particulares

Cada línea de vigilancia y, a su vez, cada evento cuenta con dimensiones que describen en más detalle las características de cada lesión en particular. Por ejemplo, todas las lesiones de causa externa no fatal registran datos sobre el mecanismo u objeto de la lesión, la presencia de alcohol o drogas en la víctima, el sitio anatómico afectado y el destino del paciente. Este conjunto de variables son compatibles solo para las lesiones de tipo no fatal y se pueden reutilizar en aquellas tablas de hechos asociadas a dicho tipo de lesión.

Sin embargo, cada tipo de lesión cuenta con sus propias variables que describen únicamente cada evento en particular. Por ejemplo, las lesiones no fatales de tránsito recopilan datos sobre el tipo de transporte y elementos de seguridad del lesionado que no aplican a otro tipo de lesiones como violencia interpersonal o lesiones autoinflingidas. Dichos datos se modelaron como dimensiones particulares que se asocian solo a una de las tablas de hechos.

3.4 Métricas

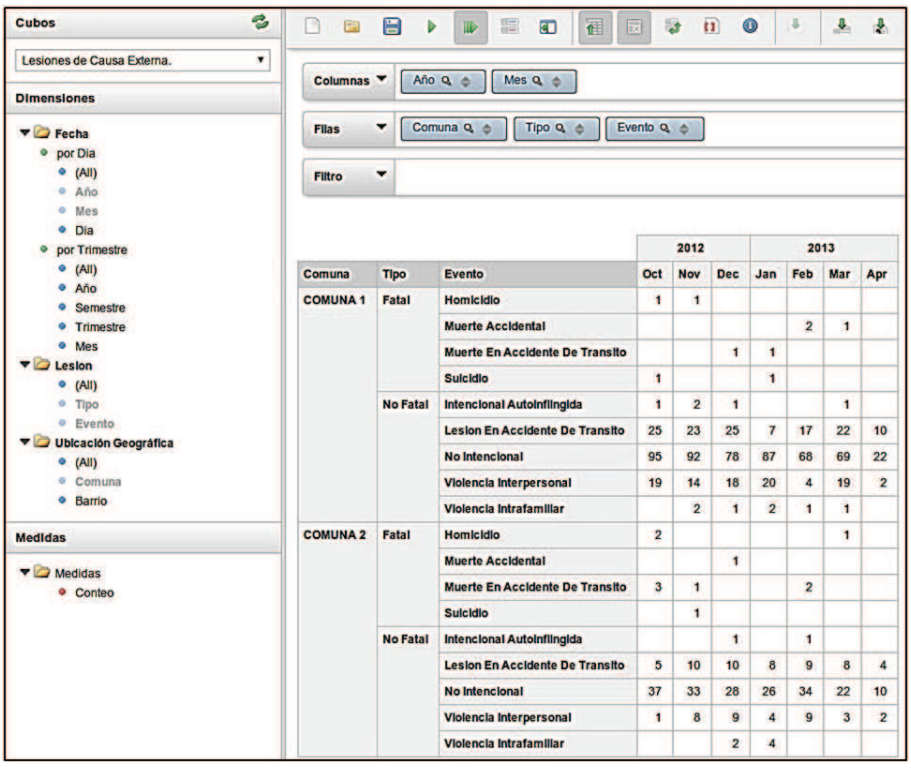
Una de las características fundamentales que vigila el Observatorio del Delito para cada uno de los eventos es el número de casos. A partir

de dicha métrica y en la implementación posterior de los cubos OLAP es posible calcular métricas adicionales como el promedio de casos y variaciones. En los prototipos implementados todas las tablas de hechos cuenta con el número de casos como métrica principal.

3.5 Construcción y visualización de cubos

Se construyeron cubos para cada tabla de hechos definida que permiten analizar información sobre homicidios, suicidios, muertes accidentales, muertes en accidentes de tránsito, lesiones interpersonales, lesiones autoinflingidas, lesiones en accidentes de tránsito y violencia intrafamiliar. Se asociaron a cada uno de ellos las dimensiones compatibles y particulares que les corresponden.

Los esquemas de los cubos OLAP construidos con PSW se publicaron en una servidor *Pentaho* y se usó el *plugin* de visualización de *Saiku Analytics* (parte de PAS) para las generación de tablas pívot dentro del servidor *Pentaho* y la visualización de los cubos. Las figuras 4 y 5 ilustran los resultados obtenidos de dicha visualización.



			2012			2013			
Comuna	Tipo	Evento	Oct	Nov	Dec	Jan	Feb	Mar	Apr
COMUNA 1	Fatal	Homicidio	1	1					
		Muerte Accidental					2	1	
		Muerte En Accidente De Transito			1	1			
		Suicidio	1			1			
	No Fatal	Intencional Autoinflingida	1	2	1			1	
		Lesion En Accidente De Transito	25	23	25	7	17	22	10
		No Intencional	95	92	78	87	68	69	22
		Violencia Interpersonal	19	14	18	20	4	19	2
		Violencia Intrafamiliar		2	1	2	1	1	
COMUNA 2	Fatal	Homicidio	2					1	
		Muerte Accidental			1				
		Muerte En Accidente De Transito	3	1			2		
		Suicidio		1					
	No Fatal	Intencional Autoinflingida			1		1		
		Lesion En Accidente De Transito	5	10	10	8	9	8	4
		No Intencional	37	33	28	26	34	22	10
		Violencia Interpersonal	1	8	9	4	9	3	2
		Violencia Intrafamiliar			2	4			

Figura 4. Visualización de cubos OLAP a través de tablas pívot

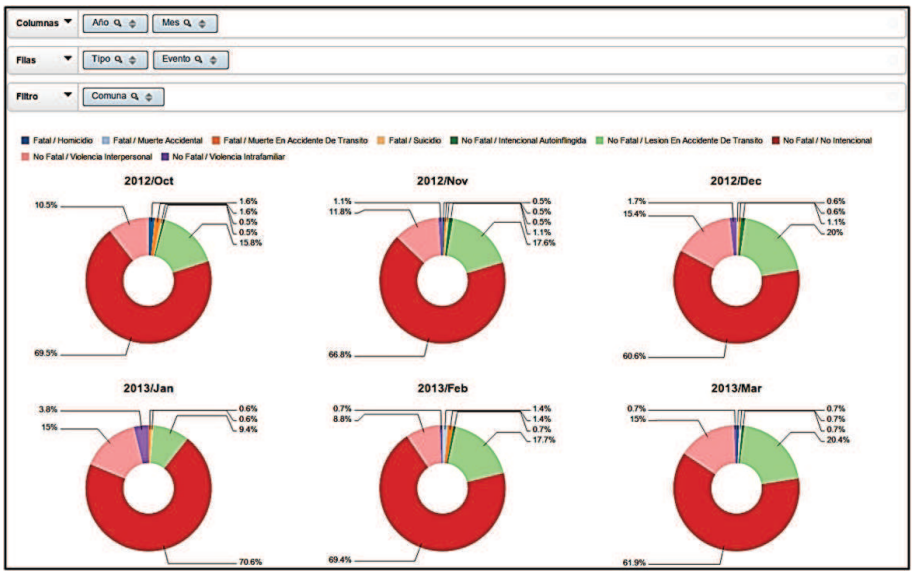


Figura 5. Visualización de cubos OLAP a través de gráficos

4. Conclusiones y trabajos futuros

Actualmente se cuenta con una implementación preliminar de un mercado de datos que soportará la detección de patrones delictivos dentro del Observatorio del Delito del municipio de Pasto. La implementación cuenta con cubos OLAP para las tres líneas de vigilancia que monitorea el Observatorio y permitirá un análisis multidimensional de todas las variables que son recopiladas por parte de dicha entidad. El nuevo almacén de datos y sus *datamarts* servirán de soporte a la toma de decisiones por parte de la administración municipal para la gestión de política más pertinentes de seguridad ciudadana y prevención del delito.

Se siguió una metodología de diseño que involucró cinco pasos: Identificación de requisitos, diseño conceptual, diseño lógico, diseño físico e implementación. Se aprovechó la interacción con el personal del Observatorio ganado durante la fase previa para avanzar en la identificación de requisitos. Los pasos de diseño se compartieron y validaron entre el grupo de desarrollo y los nuevos usuarios a través de reuniones mensuales.

Se aprovechó ampliamente la disponibilidad de herramientas de software libre presentes en el mercado para las etapas del diseño físico e implementación. Estas demostraron ser una opción viable para la construcción del almacén de datos con un costo de producción

relativamente bajo. Al ser herramientas libres se espera en el futuro incluir y acoplar los procesos ETL y visualización de cubos dentro de SIGEODEP.

Como trabajos futuros están el integrar nuevas variables a los cubos construidos y, una vez sean validados los prototipos, incluir rutinas y procedimientos de carga automática entre el sistema transaccional y el almacén de datos. Posteriormente, y como parte de la siguiente fase, se espera explotar los datos almacenados en el mercado para analizar y descubrir patrones delictivos y tendencias en los datos.

5. Agradecimientos

Este proyecto de investigación se financia con recursos del Patrimonio Autónomo Fondo Nacional de Financiamiento para la Ciencia, la Tecnología y la Innovación Francisco José de Caldas y con recursos de contrapartida de la Universidad de Nariño y de la Alcaldía Municipal de Pasto.

Referencias bibliográficas

ANAYA, R. (1996). Las bodegas de datos como apoyo a los sistemas de información acerca del negocio. En: Revista Universidad EAFIT. Vol. 32, No. 104. Medellín (Colombia): Universidad EAFIT. p 93-101. ISSN: 0120-341X.

BETANCOURT SALAZAR, C. (2005). Vigilancia de lesiones de causa externa para la toma de decisiones en el nivel local: experiencia de Pasto, Colombia, año 2005. San Juan de Pasto (Colombia): Instituto Nacional de Salud – SEA, Instituto Cisalva. <http://www.javeriana.edu.co/fcea/convocatorias/memorias_1congreso_sp/vigilancia_salud_publica/103w.pdf> [consulta: 16/03/2014]

CONNOLLY, T. & BEGG, C. (2005). Sistemas de bases de datos: un enfoque práctico para diseño, implementación y gestión. Madrid (España): Pearson Addison Wesley, 4 ed. 1269 p. ISBN: 84-7829-075-3.

ELMASRI, R. & NAVATHE, S. (2007). Fundamentos de Sistemas de Bases de Datos. Madrid (España): Pearson Addison Wesley, 5 ed. 988 p. ISBN: 978-84-7829-085-7.

HERNÁNDEZ, O.J.; RAMÍREZ, Q.M. & FERRI, R.C. (2005). Introducción a la Minería de Datos. Madrid (España): Pearson Prentice Hall. 656 p. ISBN: 84-205-4091-9.

IMHOF, C.; GALEMMO, N. & GEIGER, J. (2003): Mastering Data Warehouse Design: Relational and Dimensional Techniques. Indianapolis, Indiana (USA): Wiley Publishing Inc. 457 p. ISBN: 0 -471-32421-3.

INMON, W. H. (2000). Building the Data Warehouse: Getting Started [on line]. n.l.: Inmoncif.com. <http://inmoncif.com/inmoncif- old/www/library/whiteprs/ttbuild.pdf> [consult: 15/01/2013].

INMON, W. H. (2005). Building the Data Warehouse. Indianapolis, Indiana (USA): Wiley Publishing Inc., 4 ed. 576 p. ISBN: 978-0-7645-9944-6.

KIMBALL, R. & ROSS, M., (2002). The Data Warehouse Toolkit, the Complete Guide to Dimensional Model. New York (USA): Wiley & Sons Publishing Inc., 2 ed. 447 p. ISBN: 0-471-20024-7.

OPS (2008a). Sistematización de Experiencias sobre Sistemas de Vigilancia, Observatorios o Sistemas de Información de Violencia en América Latina: Observatorios de Violencia – Mejores prácticas [en línea]. Cali (Colombia): Organización Panamericana de la Salud. Observatorios de Violencia: mejores prácticas. ISBN: 978-958-9481-99-8. <http://www.paho.org/col/index.php?option=com_docman&task=doc_download&gid=105&Itemid=> [consulta: 18/02/2014]

OPS (2008b). Guía Metodológica para la Replicación de Observatorios Municipales de Violencia. Cali (Colombia): Centro Editorial Catorse SCS. 60 p. ISBN: 978-958-8404-00-4.

PERVERSI, I.; VALENGA, F.; FERNÁNDEZ, F.; BRITOS, P. & GARCÍA-MARTÍNEZ, R. (2007). Identificación y Detección de Patrones Delictivos basada en Minería de Datos. En: IX Workshop de Investigadores en Ciencias de la Computación, WICC 2007 (03-04/05/2007), Trelew (Chubut, Argentina): Universidad Nacional de la Patagonia San Juan Bosco. Memorias de WICC 2007, p. 385-389. ISBN: 978-950-763-075-0. <http://sedici.unlp.edu.ar/bitstream/handle/10915/20389/Documento_completo.pdf?sequence=1> [consulta: 15/03/2014]

PONNIAH, P. (2010). Data Warehousing Fundamentals for it Professionals, 2 ed. Hoboquen (New Jersey, USA): Wiley & Sons Publishing Inc. 602 p. ISBN: 978-0-470-46207-2.

RAINARDI, V. (2008). Building a Data Warehouse: With Examples in SQL Server. New York (USA): Apress. 541 p. ISBN: 978-1-4302-0527-2.

TIMARÁN, R., BARÓN, A., HERNÁNDEZ, G., HIDALGO, A. & BETANCOURTH, C. (2012). SIGEODEP: Un primer paso para la Detección de Patrones Delictivos con Técnicas de Minería de Datos. En: IX Jornadas Iberoamericanas de Ingeniería del Software e Ingeniería del Conocimiento, JIISIC 2012, (28-30/11/2012). Lima (Perú): Pontifica Universidad Católica del Perú. p. 87-96. ISBN: 978-612-4057-85-4.