

Inteligencia de Negocios y Salud: Análisis de emergencias médicas por accidentes en la vía pública

Lucas Fontana, Ramiro del Villar, Joaquín Musanti, Nicole Schmidt, Cristian Bigatti, Juan Miguel Moine

Universidad Tecnológica Nacional, Facultad Regional Rosario. Zeballos 1341 - Rosario
CP 2000 - Provincia de Santa Fe – Argentina
lfontana@frro.utn.edu.ar

Resumen: Analizar y concluir sobre datos reales en el ámbito de la salud pública constituye una tarea esencial para conocer patrones de comportamiento epidemiológicos y sociales. Las técnicas de Inteligencia de Negocios (IN) contribuyen a mejorar la calidad y eficiencia en la utilización de los recursos de los centros de salud y permiten establecer una mejor comunicación entre las diferentes jurisdicciones en tiempo real. En el presente trabajo, se demuestra con un caso real y concreto cómo la utilización de técnicas de IN permite obtener información útil para el proceso de toma de decisiones. Como consecuencia, se demuestra además la importancia de la calidad de los datos ingresados en los sistemas transaccionales y su impacto en el análisis de la información.

Palabras clave: salud pública, inteligencia de negocios, sistema de emergencias médicas, almacén de datos.

1 Introducción

La Salud Digital tiene como objetivo mejorar la calidad de la información disponible mediante la integración de los datos generados por las distintas jurisdicciones, programas y subsistemas de salud. Es uno de los tres ejes definidos para la efectiva implementación de la Cobertura Universal de Salud (CUS) en Argentina [1].

La cantidad de datos que se recolectan hoy en día en los sistemas de información puede ser abrumadora. Sin embargo, con las herramientas adecuadas, estos datos pueden convertirse en conocimiento. Es el conocimiento el que enriquece no sólo a las personas, sino también a los sistemas, en este caso el sistema de salud pública.

En los últimos años se han utilizado varias metodologías para conocer las estadísticas y proyecciones acerca de diferentes indicadores en la gestión de salud pública. Las técnicas de “Inteligencia de Negocios” y “Ciencia de datos” constituyen herramientas fundamentales para la explotación de los registros almacenados, de manera tal que podamos obtener rápidas conclusiones sobre los mismos desde diversas perspectivas [2][3]. La “Inteligencia de Negocios” puede ser definida como “el conjunto de metodologías, aplicaciones, prácticas y capacidades enfocadas a la creación y administración de información que permite tomar mejores decisiones a los usuarios de

una organización” [3]. En el ámbito de la salud pública Argentina, el uso de este tipo de herramientas tiene poca difusión.

Como hemos demostrado en [4], las técnicas de Inteligencia de Negocios permiten colaborar ampliamente en la toma de decisiones dentro del Sistema Integrado de Emergencias Sanitarias (SIES). En el presente trabajo, nos enfocamos en un evento particular, los “Accidentes en la Vía Pública” (AVP). Este evento es uno de los más importantes y demandados del sistema de emergencias médicas.

Cuando un operador recibe una llamada de emergencia para el servicio SIES asigna a la salida de la ambulancia un código inicial en función de la gravedad de la emergencia. Los códigos de salida pueden ser rojo, amarillo o verde, siendo el primero de ellos utilizado cuando existe riesgo de vida. Una vez efectuada la atención, el profesional determina un código “final” o “real”, el cual puede o no coincidir con el código asignado por el operador. Ambos códigos quedan registrados en el sistema transaccional para un posterior análisis. Si bien también se registra la fecha y hora de la salida, actualmente no se dispone de registros confiables acerca de la hora de llegada, dato que permitiría estimar el tiempo que tarda el móvil en brindar la asistencia.

A partir de este proceso, se destaca la importancia y el impacto que tiene la carga de datos en los sistemas transaccionales. Como se demostró en [4] la carga errónea de los datos impacta gravemente en el análisis estadístico, logrando que muchos registros sean desestimados, perdiendo una porción importante de la información que ayudaría a obtener mejores conclusiones.

El caso de estudio corresponde al Sistema Integrado de Emergencias Sanitarias (SIES) de un Municipio de la Provincia de Santa Fe. Como primera etapa dentro del estudio de los AVP, y como objetivo de este trabajo, abordamos el análisis exploratorio de los datos para comprender cómo es la demanda del servicio y la influencia que tienen diferentes variables sobre la misma.

1.1 Objetivo

El objetivo de la presente investigación es demostrar cómo las técnicas avanzadas de análisis de datos y de Inteligencia de Negocios colaboran en el ámbito de la salud pública generando conocimiento y facilitando el proceso de toma de decisiones.

2 Metodología

En esta investigación se ha utilizado un subconjunto de datos obtenido a partir del Data Warehouse desarrollado en un trabajo previo [5], donde se han integrado las bases de datos correspondientes a 4 centros de salud.

Al momento de analizar la información identificamos diferentes problemáticas en los datos. Algunos de los eventos registrados por el operador no poseen la identificación del paciente en cuestión, ya que muchas veces el paciente no está cargado en la base de datos y al momento de cargar el AVP, se omite esta información.

Como existen accidentes que involucran a múltiples pacientes, se procesaron los registros actuales para poder obtener la cantidad de eventos (salidas de ambulancia) y

la cantidad de pacientes afectados al mismo, identificándose unívocamente según la dirección, fecha y la hora.

El conjunto de datos procesado contiene 1344 registros correspondientes a salidas del SIES para los eventos del tipo AVP, entre el 01/01/2014 y el 31/12/2018, de los cuales 450 de ellos se encuentran geolocalizados. La información suministrada para este trabajo fue totalmente anónima, ya que la Secretaría de Salud adoptó todas las medidas necesarias para mantener la confidencialidad del paciente. La información disponible se detalla en la Tabla 1.

Tabla 1. Información disponible de AVP.

Atributo	Tipo de dato	Descripción	Ejemplos
ID	Entero	Identificador único de cada evento o salida	1115
Fecha y hora de salida	Fecha hora	Fecha y hora de la salida de la ambulancia	01/05/2018 14:25:31
Dirección	Texto (libre)	Dirección	“Santiago 1447”, “Ruta”, “San Salvador y Paraguay”
Motivo	Texto (libre)	Motivo de la urgencia	“Caída en la vía pública”, “Choque Auto Moto”, “AVP”
Móvil	Entero	Número de móvil que acude al lugar del hecho	1, 2
Código salida	Texto (Codificado)	Código inicial asignado por el operador.	Rojo, Verde
Código final	Texto (Codificado)	Código real o final, registrado por el profesional que acude a la urgencia.	Rojo, Amarillo
Latitud	Coordenada	Latitud	-29.154
Longitud	Coordenada	Longitud	-60.770
Hubo traslado	Booleano	Indica si la ambulancia realizó al menos un traslado a un centro asistencial.	Si
Cantidad Pacientes	Entero	Cantidad de pacientes asistidos por el evento o urgencia.	1

El Motivo también es considerado de interés ya que nos permite identificar la causa del AVP. Sin embargo, el mismo no está codificado dentro del sistema transaccional, por lo cual cada operador lo carga en un campo de texto libre. Por ejemplo: podemos encontrar motivos con las siguientes descripciones: “Auto-Moto”, “auto moto”, “auto y moto”, “avp”, etc. Esto implica que resulta compleja la exploración de la información a través de este atributo, además de la imposibilidad de identificar si la ausencia de una palabra indica que realmente no existió. Sin embargo, mediante la aplicación

de técnicas de limpieza de datos hemos podido establecer 5 motivos frecuentes, los cuales fueron tipificados en las categorías: “auto-auto”, “auto-moto”, “Caída de moto”, “Bici-Auto”, “Vuelco”.

Se agregaron también variables calculadas a partir de las ya existentes, por ejemplo la estación del año y el día de la semana. También se calculó la variable “Rango horario”, que agrupa la hora de los eventos en diferentes valores discretos (Tabla 2). Estos rangos fueron definidos de manera tal que nos permita analizar horarios críticos y compararlos entre sí.

Tabla2.Rangos Horarios.

Horario	Hora desde	Hora hasta
Madrugada	00:00:00	06:00:00
Mañana	06:00:00	12:00:00
Mediodía	12:00:00	14:00:00
Tarde	14:00:00	20:00:00
Noche	20:00:00	23:59:59

Con el objetivo de presentar de una manera clara y concisa la información alojada en el Data Warehouse, se determinó que la herramienta a utilizar sería Power BI Desktop de Microsoft [6]. Esta herramienta ofrece una versión gratuita y posee todas las posibilidades de análisis detalladas en este trabajo, permitiendo que su implementación en el sector público sea factible. Este tipo de tecnología permite al usuario interactuar con los datos utilizando filtros y visualizaciones dinámicas, de manera tal que pueda realizar un análisis integral obteniendo sus propias conclusiones [7].

3 Resultados

Para la visualización principal se decidió crear un tablero de control [8] compuesto por 4 cuadrantes como puede verse en la Figura 1.

El *primer cuadrante* muestra todos los eventos AVP que han sido geolocalizados dentro del Municipio. El tamaño de la burbuja representa la cantidad de eventos en ese punto geográfico. A la derecha del mapa encontramos un semáforo, el cual permite filtrar por la gravedad del accidente (código inicial Verde, Amarillo o Rojo). Mediante la aplicación de este filtro, rápidamente se pueden identificar zonas de mayor o menor riesgo.

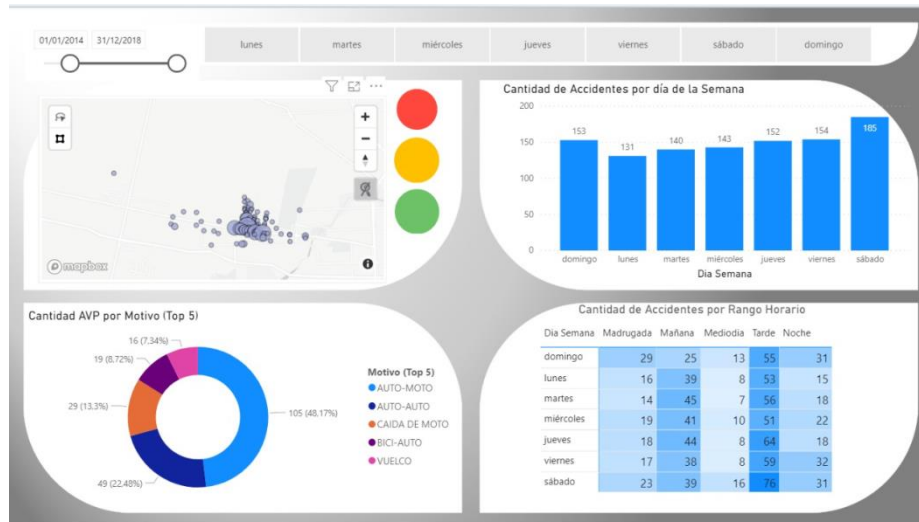


Fig. 1. Tablero de control.

En la *segundo cuadrante* encontramos un gráfico de barras el cual indica la cantidad de eventos ocurridos para cada día de la semana, en función del periodo seleccionado.

En la *tercer cuadrante* visualizamos información acerca de la variable “motivo”. Si bien es un campo complejo de analizar debido a que actualmente es texto libre, logramos tipificar algunos de ellos y evaluar su incidencia mediante un gráfico circular. Se puede observar que el motivo más frecuente es el “AVP auto-moto”. Aquí vemos la importancia de realizar cambios en los registros de este dato, ya que aporta información muy valiosa para conocer las causas de los AVP en el Municipio.

El *cuarto cuadrante* está formado por una tabla que totaliza la cantidad de eventos por día de la semana y horario. Se puede ver rápidamente que la tarde es el momento del día más propenso a tener accidentes, mientras que el pico más alto se encuentra el día sábado en este horario. Por la noche/madrugada se observa que en los fines de semana la cantidad de eventos aumentan frente al mismo horario en días hábiles. Por el contrario, durante la mañana la cantidad de eventos disminuye en los días sábados y domingos respecto a los días de semana.

Complementariamente al tablero de control, se realizaron diversas gráficas multivariantes cruzando toda la información disponible.

En la Figura 2 se observa que la cantidad de salidas de SIES por AVP se mantiene prácticamente constante a lo largo del tiempo. Esta observación nos resultó de gran interés, ya que podemos visualizar que la demanda de un evento aparentemente aleatorio en realidad es altamente predecible.

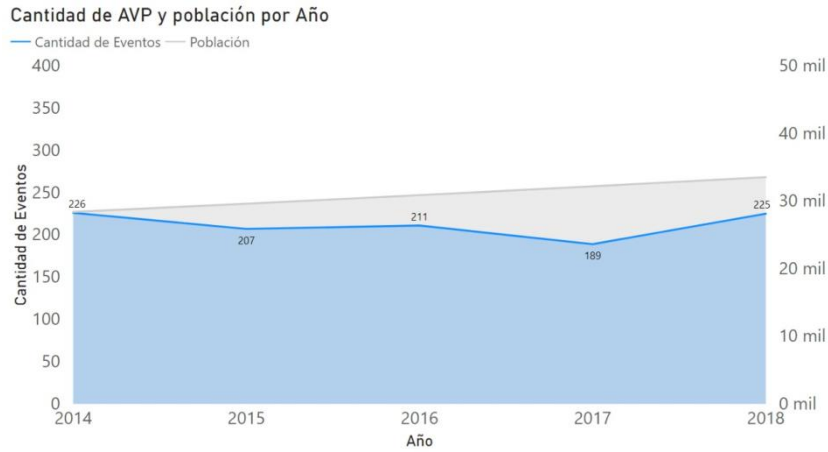


Fig. 2. Cantidad de AVP y población del Municipio por año.

Además, como se representa en la Figura 2, si bien la cantidad de AVP se ha mantenido constante, según el Instituto Provincial de Estadísticas y Censos de Santa Fe [9] la cantidad de habitantes aumentó considerablemente, como así también la cantidad de vehículos que transitan en la zona, lo que podría indicar una reducción de la cantidad de accidentes por vehículo. Podemos hacer una comparación de este indicador con [10] en donde se demuestra una caída en la tasa de mortalidad desde el año 2014 en adelante. Para visualizar mejor esta información, construimos un indicador que representa la cantidad de AVP anual cada mil habitantes. Como podemos ver en la Figura 3, partiendo de 8 AVP anuales/1000 hab. en 2014 se evidencia una reducción hacia 2018 de 6,7 AVP anuales/1000 hab.

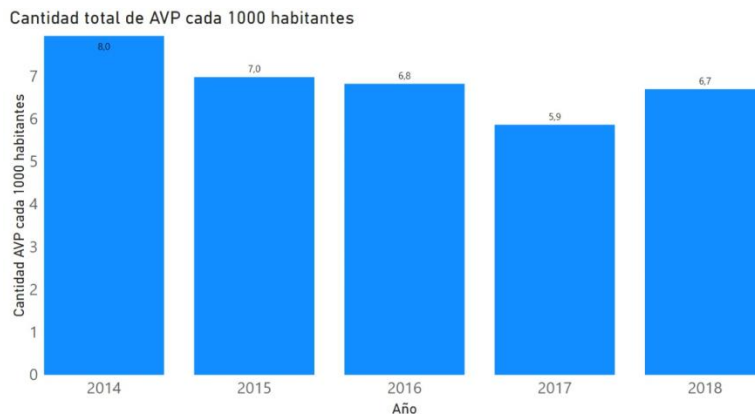


Fig.3. Cantidad de AVP anual cada 1000 habitantes

En otro análisis multivariante se estudió el comportamiento de la variable “Código de Salida”. En concreto, se puede apreciar en la Figura 4 que la gravedad del evento varía dependiendo del día de la semana. De todos los eventos ocurridos en los días lunes, el 83,2% corresponde a llamadas de emergencia en donde existe riesgo de vida, siendo este día donde se ve la mayor proporción de salidas con código “rojo”.

Si con el mismo enfoque analizamos y filtramos los datos sobre los accidentes ocurridos en “verano”, en la Figura 4 vemos que en general hay un aumento en la proporción de eventos con código “rojo”, especialmente el día domingo (40,3% durante todo el año en comparación de 58% en verano).

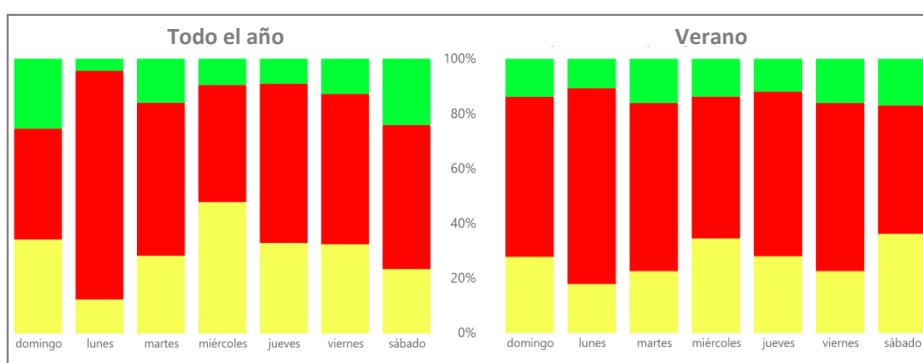


Fig. 4. Porcentajes de AVP según el código de salida y día de la semana, teniendo en cuenta todos los meses del año (izquierda) y filtrando solo meses de verano (derecha).

Por último se analizaron las salidas de acuerdo al momento del día y a la estación del año (Figura 5). Aquí se observan dos particularidades. La primera, como se mostró anteriormente, la “tarde” es el momento del día con mayor proporción de eventos. La segunda es que en verano se visualiza en general una mayor proporción de eventos, especialmente en el horario nocturno. En esta estación, el Municipio triplica la cantidad de habitantes debido a su abundancia de “casas de veraneo”, lo cual podría explicar la gran proporción de eventos.

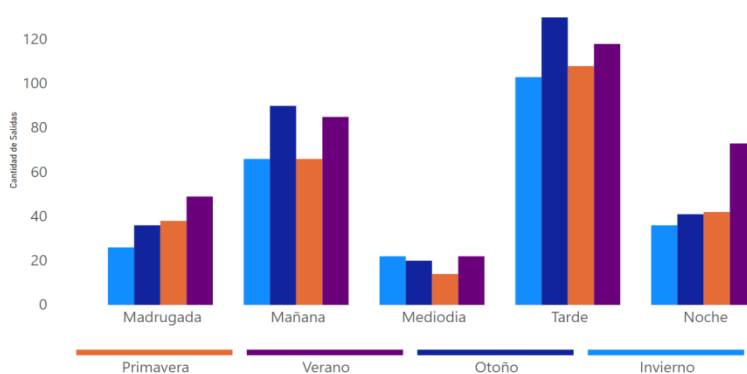


Fig. 5. Cantidad de eventos AVP por rango horario y por estación del año.

4 Conclusiones

En el presente trabajo se ha expuesto cómo las técnicas de Inteligencia de Negocios y explotación de datos pueden colaborar a la toma de decisiones en el ámbito de la salud pública basándose en datos reales.

Como resultado, se pudo analizar en forma dinámica y gráfica los datos disponibles del servicio de emergencias SIES para los eventos AVP, observando particularmente en este caso una demanda casi constante, incluso cuando la cantidad de vehículos y personas que transitan por la zona es creciente.

También se visualizaron claramente los momentos del día en donde se tiene una mayor demanda del servicio, información de utilidad para determinar, por ejemplo, en qué días y horarios sería más conveniente realizar mantenimiento a los vehículos y cuándo se requerirá de una mayor atención y disponibilidad por parte del personal.

Particularmente en este caso de estudio sería importante replantear el registro del motivo del AVP, el cual debería estar clasificado en categorías predeterminadas para aportar mayor información sobre las causas de los accidentes en el Municipio.

A su vez, se ha podido evidenciar la importancia del proceso de carga de datos, que nos lleva a concientizar acerca del valor de la información presente en los sistemas transaccionales. Datos bien cargados conducen a análisis más exactos, confiables y de los cuales se pueden extraer importantes conclusiones, obteniendo un estudio de mayor calidad.

Al utilizar estas herramientas es posible mejorar la calidad y eficiencia en la utilización de los recursos de los centros de salud. Permiten establecer una mejor comunicación entre las diferentes jurisdicciones y sistemas tanto nacionales como provinciales, disminuyendo brechas de calidad en la atención de la salud utilizando herramientas innovadoras.

Este trabajo fue realizado en el marco del Proyecto de Investigación y Desarrollo: "Big Data y Salud: Un análisis interdisciplinar para la construcción de insumos en el marco de políticas públicas para el desarrollo". PID UTN FA 5093.

5 Trabajo futuro

Como segunda etapa de este trabajo nos centraremos en la generación de modelos más complejos con técnicas de Minería de datos, que permitan encontrar patrones de comportamiento predictivos y descriptivos.

Además, se incluirá en este estudio información correspondiente a servicios de emergencias de otras localidades para obtener un mayor volumen de datos y nuevas variables.

Referencias

1. Ministerio de Salud y Desarrollo Social: Estrategia Nacional de Salud Digital, https://www.argentina.gob.ar/sites/default/files/estrategia-nacional-de-salud-digital_resolucion_189_2018.pdf. Último acceso 18/06/2020.

2. Mettler, T., & Vimarlund, V.: Understanding business intelligence in the context of healthcare. *Health informatics journal*, 15(3), 254-264 (2009).
3. Díaz, J. C.: *Introducción al Business Intelligence*. Editorial UOC. (2009)
4. Fontana, L., Del Villar, R., Musanti, J., Schimdt, N., Angeloni, L., Boffeli, D., Bigatti, C., Moine, J. M.: Análisis de emergencias médicas con técnicas de Inteligencia de Negocios. *AJEA*, (1). (2019)
5. Fontana, L., Del Villar, R., Musanti, J., Schimdt, N., Angeloni, L., Boffeli, D., Bigatti, C., Moine, J. M.: Aplicación de técnicas de Business Intelligence en el ámbito de la salud pública. *Revista "Rumbos Tecnológicos"*, Vol 11, año ISSN 1852-7701. (2019)
6. Ferrari, A., & Russo, M.: *Introducing Microsoft Power BI*. Microsoft Press. (2016)
7. Microsoft Corporation. *Microsoft Power BI Desktop*. <https://powerbi.microsoft.com>
8. Ballvé, A.: Creando conocimiento en las organizaciones con el Cuadro de Mando Integral y el Tablero de Control. *Revista de Contabilidad y Dirección*, 3, 13-38. (2006)
9. Instituto Provincial de Estadísticas y Censos de la Provincia de Santa Fe: Proyección y Estimación de la población, <http://www.estadisticasantafe.gob.ar/contenido/proyeccion-y-estimacion-de-la-poblacion>. Último acceso el 18/06/2020.
10. Ministerio de Salud, Dirección de Estadística e Información de Salud: Mortalidad por lesiones en accidentes de tránsito, <http://www.deis.msal.gov.ar/wp-content/uploads/2019/02/SintesisNro4.pdf>. Último acceso 18/06/2020.