

# Caracterización de los registros en emergencias atendidas por INPRADEM utilizando minería de datos

*Characterization of records in emergencies assisted by INPRADEM using data mining*

Yamilet Casanova\*, Francisco Hidrobo\*\* y Lucileima Rosales\*\*\*

Código JEL: C44, C89, D83

Recibido: 25/03/2019, Revisado: 21/05/2019, Aceptado: 25/06/2019

## Resumen

En este artículo se presenta la caracterización de los registros de emergencias de INPRADEM utilizando minería de datos. Se usó la metodología CRISP-DM, haciendo énfasis en: emergencias suscitadas en el estado Mérida, colisiones por horas, semanas con mayor cantidad de emergencias y eventos en períodos vacacionales. Los resultados aportan información relevante relacionada a los equipos más usados, distribución de colisiones por horas, emergencias que requieren o no equipos, comportamientos en períodos vacacionales y días de la semana con mayor número de eventos. INPRADEM cuenta con una nueva forma de ver sus datos, prometiendo beneficios a la solución de una gran variedad de problemas como: planeación económica, distribución de trabajo y equipo, análisis de servicio y prevención en épocas de demanda alta.

**Palabras claves:** organización, minería de datos, caracterización.

## Abstract

This article presents the characterization of INPRADEM's emergency records by using data mining. The methodology used was CRISP-DM, placing an emphasis on: emergencies caused in the Merida State, collisions of hours, weeks with more emergencies and events during vacation periods. The results provide relevant information related to the most used equipment, collision arrangement by hours, emergencies that do or do not require equipment, performance in holiday periods, and days of the week with a greater number of events. INPRADEM has a new way of viewing its data, promising benefits to the solution of a wide range of problems such as: economic planning, work and equipment distribution, service analysis and prevention in periods of high demand.

**Key words:** Organization, data mining, characterization

\* Ingeniero de Sistemas de la Universidad de Los Andes. Mérida, Venezuela. Teléfono: (58) 04126630203. Correo electrónico: yamiccv@gmail.com.

\*\* Ph.D. En Arquitectura de Computadores de la Universidad Politécnica de Cataluña. Escuela de Ciencias Matemáticas y Computacionales Yachay Tech. Ecuador. Teléfono: (593) 980610529. Correo electrónico: fhidrobo@yachaytech.edu.ce, fhidrobo@gmail.com.

\*\*\* Magister Scientiae en Estadística de la Universidad de Los Andes. Escuela de Estadística. Universidad de Los Andes. Mérida, Venezuela. Teléfono: (58) 04147282629. Correo electrónico: lucileima@ula.ve, lucileimarosales@gmail.com.

## 1. Introducción

Durante los últimos años el crecimiento de las zonas urbanas ha traído consigo el aumento en las necesidades de movilidad, llevando a un incremento del parque automotor, que para el año 2013 llegó a 102.925 carros y motos en el Estado Mérida según el Instituto Nacional de Estadística (INE). A la cantidad de vehículos se suma la falta de nuevas avenidas, planes de movilidad y otras iniciativas como campañas de educación a conductores que contribuyen a un elevado número de accidentes de tránsito, cifra que aumenta junto con la población y cantidad del parque automotor.

Mérida, por su relieve montañoso y siendo zona sísmica, presenta lugares de alto riesgo que son utilizados para desarrollos urbanos, atentando contra la integridad física de las personas que allí habitan. Por otra parte, Mérida es una ciudad turística, y con su creciente población tiene muchas actividades (concentraciones, manifestaciones, conciertos) y eventos públicos (i.e. las ferias del sol) que requieren apoyo de instituciones que presten servicios de atención en caso de presentarse contratiempos.

En vista de todas estas necesidades que presenta el estado, y de manera particular la ciudad, la Gobernación del Estado Mérida el 03 de octubre de 1996 creó la Fundación para el Manejo de Emergencias, Desastres Naturales y Defensa Civil del Estado Mérida (FUNDEM), que evolucionó a través de los años y se convirtió en el Instituto de Protección Civil y Administración de Desastres del Estado Mérida (INPRADEM), cuya misión es:

Ejecutar acciones que implican la gestión de riesgo y minimización de los efectos de los desastres en cada una de sus etapas, preparación, prevención, mitigación, alerta, respuesta, rehabilitación y recuperación, con un talento humano proactivo y profesional sustentado en los valores de voluntariedad, solidaridad, sentido social y humanitario (Emergencia 171 Mérida, 2018-2019, parr. 7).

En aras de mejorar los procesos internos de la organización, INPRADEM implementó en el año 2013 una planilla como mecanismo para llevar el registro de las emergencias que atienden en los diversos sectores del Estado. Dicha planilla contiene información sobre unidades despachadas, emergencias atendidas (nombre de los oficiales que atendieron dicha emergencia, tipo de emergencia, hora de despacho y llegada de la unidad), materiales usados en el servicio prestado o emergencia atendida y en algunos casos, los nombres de las personas que fueron atendidas. Además, se registra si se trabajó en conjunto con otro cuerpo de seguridad. Esta planilla es digitalizada mensualmente en un documento en formato Excel para tener la información tanto en físico como en digital. Sin embargo, el tiempo y esfuerzo empleado por los trabajadores para llevar el registro de todos los servicios, no se está aprovechando. Con el fin de contribuir con este organismo que trabaja en pro de la población del Estado Mérida, y tomando en cuenta que los datos recopilados en la planilla no han sido analizados, se propone caracterizar dichos datos mediante técnicas de minería de datos, para dar soporte a la toma de decisiones en los procesos de gestión de la organización.

En tal sentido, dichos datos pueden ser analizados para aportar información útil que contribuya a la mejora de la gestión de esta organización mediante la minería de datos, que según Zabala es una técnica que consiste:

[En la] extracción no trivial de información que reside de manera implícita en los datos. Dicha información era previamente desconocida y podrá resultar útil para algún proceso. En otras palabras, la minería de datos prepara, sondea y explora los datos para sacar la información oculta en ellos (2014, p.6).

Además de esta introducción, el artículo contiene otras cinco secciones: marco de referencia, metodología, fases de la investigación, resultados y, por último, se presentan las conclusiones.

## 2. Marco de referencia

A continuación, se presentan dos subsecciones, la primera contiene una descripción de la organización estudiada (INPRADEM) y en la segunda, se mencionan algunos trabajos relacionados con la investigación.

### 2.1. Origen de INPRADEM.

Dadas las condiciones geográficas y geomorfológicas que posee el Estado Mérida, la Gobernación de esta entidad creó la Fundación para el Manejo de Emergencias, Desastres Naturales y Defensa Civil del Estado Mérida (FUNDEM) el 03 de octubre de 1996 con el objeto de ofrecer atención especializada en materia de emergencias. Dicha Fundación recibe el nombre de Instituto el 03 de febrero de 1999 según Decreto Oficial N° 106 Extraordinaria, emanado del Despacho de la Asamblea Legislativa del Estado Mérida; asumiendo a partir de esta fecha las funciones para la atención de emergencias, desastres naturales y defensa civil del Estado Mérida (Emergencia 171 Mérida, 2018-2019). De esta manera, las personas responsables de su creación diseñaron un conjunto de programas y estrategias de atención y prevención a favor de la población de Mérida.

En tal sentido, ante la situación de riesgo natural y alta incidencia de accidentes antrópicos que distinguen al Estado Mérida, se crea el Instituto de Protección Civil y Administración de Desastres del Estado Mérida (INPRADEM) el 01 de octubre de 2001. Dicho organismo es de carácter público, social y humanitario y tiene como propósito resguardar a la persona y a la sociedad en general ante las eventualidades provocadas por agentes naturales o humanos, así como disminuir los efectos negativos de tales desastres.

### 2.2. Revisión de la literatura

Sobre esto, Aránguiz (2012), realizó una prueba basada en la aplicación de técnicas de minería de datos de segmentación, asociación y clasificación, con el objeto de identificar patrones y variables influyentes en los accidentes de tránsito. Adicionalmente, efectuó un

análisis difuso del problema, mediante la creación de atributos difusos, que aportaran mayor información sobre los accidentes, acompañados de algoritmos difusos, particularmente de clasificación. Para dicha prueba, el autor tomó datos de las principales zonas del país durante el período 2007 - 2009. El resultado de esta investigación fue un plan de prevención de accidentes de tránsito.

Por su parte, Hassinger (2015) utilizó árboles de decisión para analizar los accidentes de tránsito, identificando las variables con mayor relevancia en la gravedad del accidente y extrayendo reglas de decisión para descubrir patrones que sirvan a los analistas, y gestores de seguridad vial, para realizar planes concretos con el fin de reducir el impacto socioeconómico causado por dichos accidentes.

Así, Salinas y Vele (2014), realizaron una predicción de la tendencia de la cantidad de accidentes de tránsito registrados desde agosto de 2014 hasta 2016. Posteriormente, mediante indicadores de accidentalidad de tránsito, abordaron un análisis descriptivo. De igual manera, los autores desarrollaron un análisis de series temporales utilizando el modelo autorregresivo integrado de medias móviles (ARIMA) para obtener la proyección sobre el aumento de los accidentes viales. En función de los resultados obtenidos en los análisis plantearon campañas y reformas orientadas en la disminución de accidentes de tránsito en el cantón Cuenca.

Por otro lado, García, Delgado, Díaz y García (2012), realizaron un análisis de las carreteras rurales de la provincia de Villa Clara ubicada en la región central de Cuba para estudiar la seguridad vial de los tramos que conforman dicha red. Para ello, los autores clasificaron los accidentes vehiculares de acuerdo a cinco aspectos y determinaron la influencia de cada uno de los componentes del sistema de seguridad vial en la accidentalidad y su interrelación. Los resultados obtenidos permitieron conocer las causas que ocasionaban dichos accidentes y en función de eso, los autores propusieron medidas para la disminución de los mismos.

Por su parte, Villarino (2015) aplicó minería de datos para el estudio de tablas de siniestralidad vial, usando modelos de clasificación basados en regresión logística. Por otra parte, utilizó

diversos métodos de ensamble de clasificadores mediante la técnica de stacking con el objeto de mejorar la precisión obtenida con los modelos de clasificación y reducir la varianza de los errores cometidos, observándose mejoras en cuanto a la precisión aportada por las técnicas de minería de datos frente al modelo clásico de regresión logística. Para la prueba se consideraron los 285.919 contenidos en la base de datos de accidentes ocurridos en España durante el año 2012 proporcionada por la Dirección General de Tráfico, de los cuales, 83.115 correspondían a registros de accidentes con víctimas por los diferentes cuerpos de policía y guardia civil y 202.804 registros referentes a datos específicos de las propias víctimas.

### 3. Metodología

La minería de datos es un dominio de la ciencia de la computación que permite el análisis de grandes cantidades de datos para encontrar y extraer patrones significativos útiles para el proceso de la toma de decisiones<sup>1</sup>. Existen diversas definiciones, por ejemplo, para Govindarajan y Chandrasekaran (2011) es el uso de algoritmos para extraer la información y patrones derivados por el proceso de descubrimiento de conocimiento en bases de datos. Por su parte, Larose y Larose (2014) la define como el proceso de descubrir nuevas correlaciones, patrones y tendencias significativas a través de grandes cantidades de datos, utilizando técnicas estadísticas, matemáticas y reconocimiento de patrones. Existen diversas metodologías para abordar el análisis de grandes cantidades de datos. En tal sentido, León (2017) destaca: SEMMA, KDD y CRISP-DM. La metodología usada para el desarrollo del proyecto fue CRISP-DM (*Cross-Industry Standard Process for Data Mining*) y el tratamiento de los datos se realizó en Excel de Microsoft Office 2013 y RStudio.

#### 3.1. Metodología CRISP-DM:

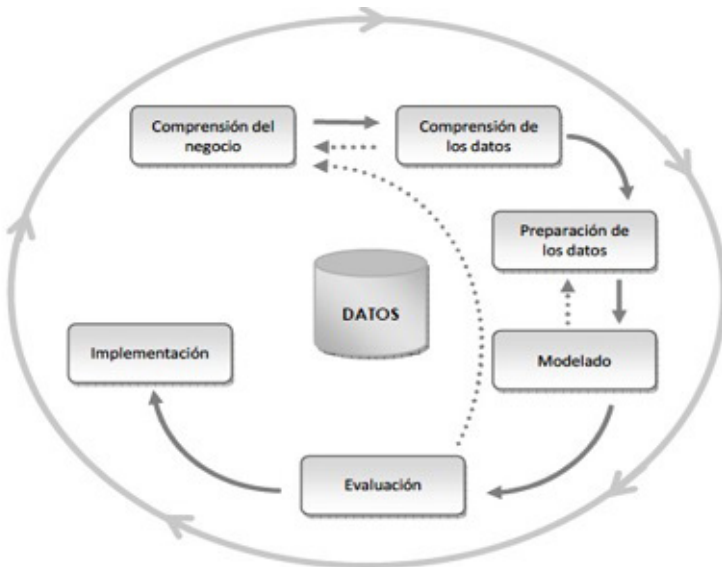
##### *Generalidades*

El acrónimo CRISP-DM (*Cross- Industry Standard Process for Data Mining*) fue construido en 1993 por líderes de un grupo de industrias

como Daimler Benz, SPSS de Inglaterra, OHRA de Holanda, NCR de Dinamarca, consorcio de empresas Europeas y AG de Alemania, con el objeto de proveer aportes en el área de minería de datos (Timarán, Hernández, Caicedo, Hidalgo y Alvarado, 2016). La metodología CRISP-DM “se describe en términos de un modelo de procesos jerárquicos donde las tareas están organizadas en cuatro niveles de abstracción (de general a específico): fases, tareas generales, tareas especializadas e instancias de proceso” (Chapman, Clinton, Kerber, Khabaza, Reinartz, Shearer y Wirth, 2000, p. 8).

*Ciclo de vida de la metodología CRISP-DM*

El ciclo de vida de un proyecto de minería de datos es dinámico e iterativo e implica las siguientes fases: comprensión del negocio o problema, comprensión de los datos, preparación de los datos, modelado, evaluación e implementación (ver figura 1).



**Figura 1.** Fases del proceso CRISP-DM.

Fuente: Chapman et al. (2000).

Cada fase está constituida por tareas generales, y cada una de éstas se desglosan en tareas específicas donde se describen las acciones a desarrollar para determinadas situaciones (Larose y Larose, 2014). En la figura 2 se presenta una visión general de las tareas correspondientes a cada una de las fases.



Figura 2. Tareas de la metodología CRISP-DM.

Fuente: Elaboración propia.

Actualmente, CRISP-DM es una de las metodologías más usadas en minería de datos por su flexibilidad en la secuenciación de sus fases.

#### 4. Fases de la investigación

En esta sección se describen cada una de las fases del proyecto, presentándose para cada prueba solo una parte de los resultados obtenidos<sup>2</sup>.



#### 4.1. Fase 1: Comprensión y análisis del problema

Para dar soporte a la toma de decisiones en los procesos de gestión de INPRADEM, se analizan los datos registrados de las emergencias atendidas por dicha organización mediante técnicas de minería de datos.

#### 4.2. Fase 2: Comprensión de los datos

Los datos proporcionados por la organización corresponden al período 2014 - 2016 de todos los centros de trabajo ubicados en el Estado Mérida, con un total de ocho (08) centros que trabajan en conjunto para satisfacer la demanda del Estado. Dichos datos se encuentran en una planilla que registra las actividades realizadas por INPRADEM y están digitalizadas en Excel de Microsoft Office. La planilla es manejada en cada centro y enviada a la central del Estado ubicada en la ciudad de Mérida. Siendo así, por cada filial que tiene INPRADEM en el Estado se tiene un documento al mes. La información contenida detalla la actividad que realiza el o los funcionarios despachados para atender dicha actividad. En el cuadro 1 se presenta una breve descripción de los datos.

**Cuadro 1.** Variables con su descripción.

Nombre del campo	Tipo de campo	Descripción
Semana	Texto	Semana del año en la que ocurre el evento
Fecha	Fecha	Fecha exacta del evento definida: día de la semana, día del mes, mes y año. Ejemplo: sábado, 2 de abril de 2016
Hora de inicio	Hora	Hora en que es despachada una unidad para el evento
Hora final	Hora	Hora en que la unidad regresa o culmina el evento
Total horas efectivas	Hora	Duración de la atención al evento
Cantidad de funcionarios	Número entero	Cantidad de funcionarios que se dispusieron para el evento
Horas hombre	Hora	Cantidad en horas efectivas por cantidad de funcionarios que prestaron el servicio.
Número de producto meta	Texto	Código empleado para identificar la clase de servicio que se prestó para el evento. Ejemplo: Producto_Meta_7
Denominación producto meta	Texto	Especificación del servicio prestado. Ampliando el código del producto meta y describiendo la actividad con detalle.
Actividad	Texto	Descripción general del producto. Ejemplo: Eliminación de riesgo. Traslado de emergencia

Fuente: Elaboración propia.

### 4.3. Fase 3: Preparación de los datos

En esta fase se realizó una limpieza de los datos, seleccionándose para el estudio un total de once (11) variables: semana, día, hora de inicio, hora final, total de horas efectivas, cantidad de funcionarios, número de producto meta, actividad, municipio, departamento, equipos y materiales utilizados. Dichas variables se clasificaron en tres grandes grupos: (1) Aquellas que están relacionadas con el tiempo; (2) La variable actividad, que no tiene relación alguna con el tiempo ni el espacio geográfico en que se desarrolle, pero ella si influye en otras variables y (3) Espacio geográfico, conformado por todas las variables que dependen del lugar en que se desarrolle la actividad. Las variables número de funcionarios y equipos dependen tanto del espacio geográfico como de la actividad debido a que los equipos que se usan dependen del servicio que presta; así como también del lugar, ya que en diferentes sectores del Estado se tiene a disposición distintas cantidades de funcionarios y equipos. En la figura 3 se presenta el esquema de clasificación considerado para las variables.

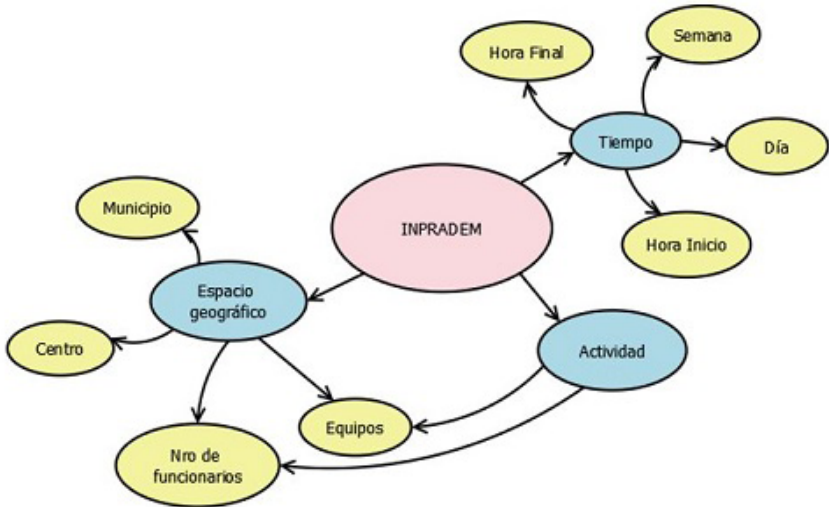


Figura 3. Esquema de clasificación de las variables.  
Fuente: Elaboración propia.

#### 4.4. Fase 4: Modelación

##### 4.4.1. Pruebas Iniciales

##### Gráficos de Frecuencias

Se realizó una primera prueba basada en el cálculo de frecuencias máximas por variable (ver figura 4).

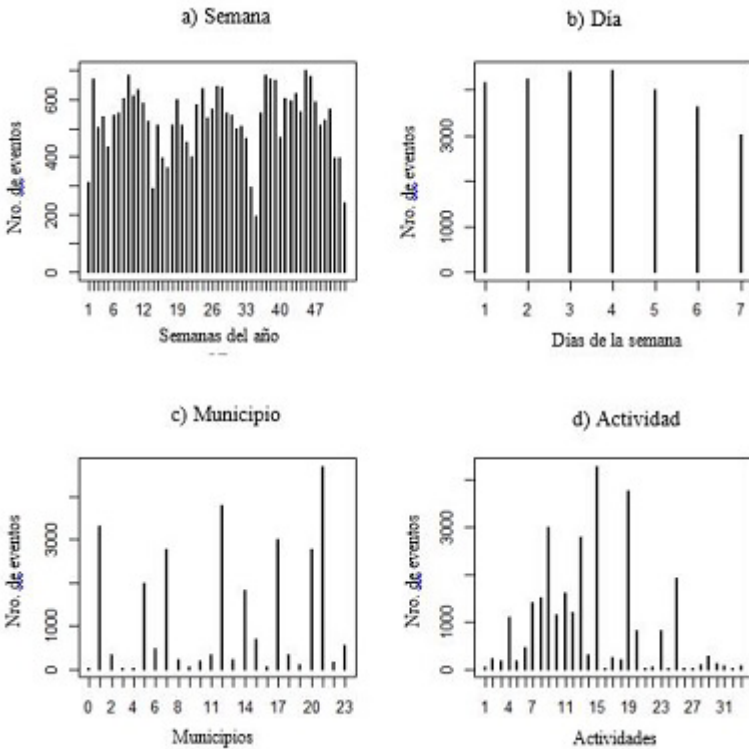


Figura 4. Gráficos de frecuencias

Fuente: Elaboración propia.

##### *Por día las horas en que se presentan los eventos*

Se consideró para cada día de la semana la hora en que se despachan las unidades para atender las diferentes actividades (ver figura 5).

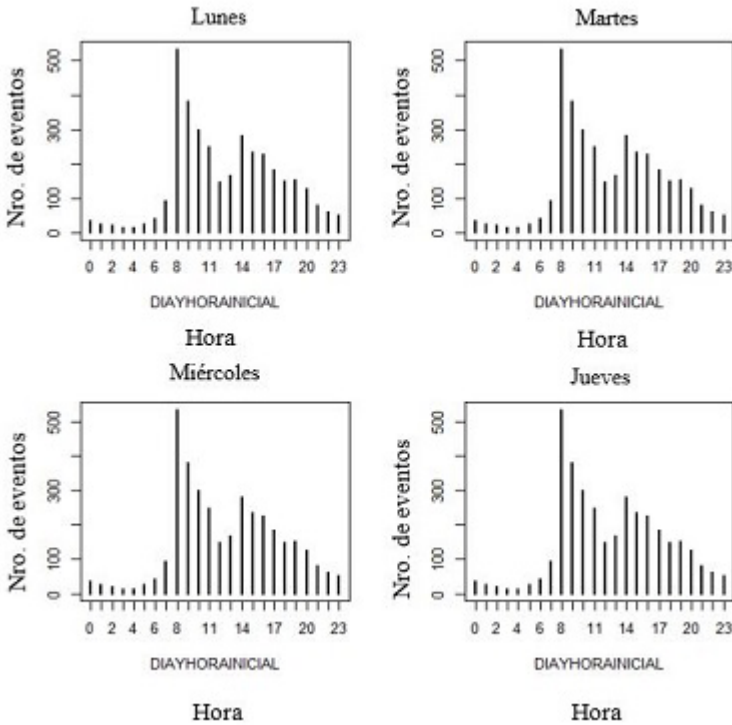


Figura 5. Horas en que se presentan los eventos por día.  
Fuente: Elaboración propia.

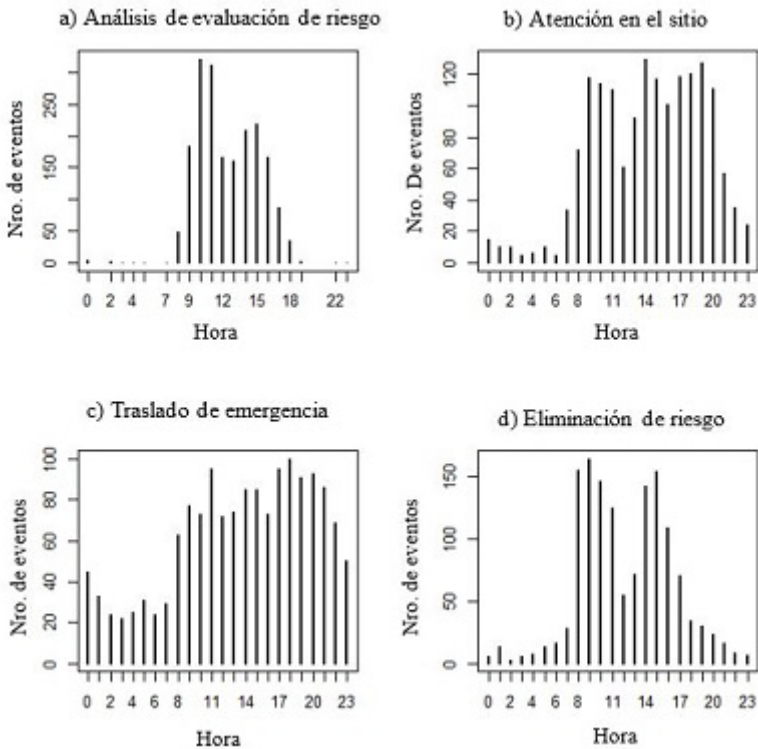
*De las actividades que más se realizan en la organización se estudia el comportamiento de las horas en que se presta el servicio*

Para esta prueba se tomaron las actividades con mayor frecuencia dentro de la organización y se observó el comportamiento de las diferentes variables con respecto a cada actividad (ver figura 6).

#### 4.4.2. Pruebas en torno a las actividades no programadas

Para esta serie de pruebas centradas en las actividades no programadas se consideró una nueva variable llamada centro. La inclusión de dicha

variable obedece a que en muchos casos, por la ubicación geográfica, una emergencia es atendida por un centro que usualmente no atiende emergencias en ese municipio, dando a entender que un centro puede atender emergencias que se suscitan en varios municipios.



**Figura 6.** Horas en que se presta el servicio por actividad.  
Fuente: Elaboración propia.

*Comportamiento general de actividades con respecto a cada variable*

En el cuadro 2 se describe el comportamiento general de cada variable con sus valores más altos, respectivamente.

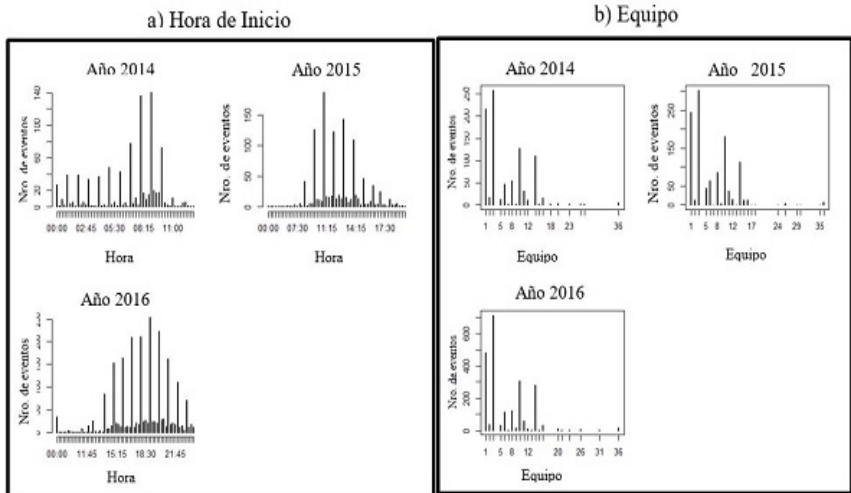
*Comportamiento por año de las actividades con respecto a cada variable*

Para esta prueba se agregó la variable año separando las emergencias suscitadas por año para analizar el comportamiento de las variables en el transcurso de los tres años bajo estudio (ver figura 7).

**Cuadro 2.** Comportamiento de las actividades no programadas por variable.

Variables	Resultado (cantidad de eventos)
Semana	Segunda semana del año (117), tercera semana de marzo (117), segunda semana de septiembre (111).
Día	Los días se comportan similarmente, no se observa ningún día que destaque en particular. Jueves (649), sábado (646), miércoles (640).
Hora de Inicio	Se presentan en un horario tarde-noche. En formato de 24 horas: 19 (323), 18 (309), 20 (301), 17 (292)
Departamento	SAR (1644), Respuesta Inmediata (580)
Centro	Sur del Lago (856), Panamericana (576) y Libertador (541)

Fuente: Elaboración propia.

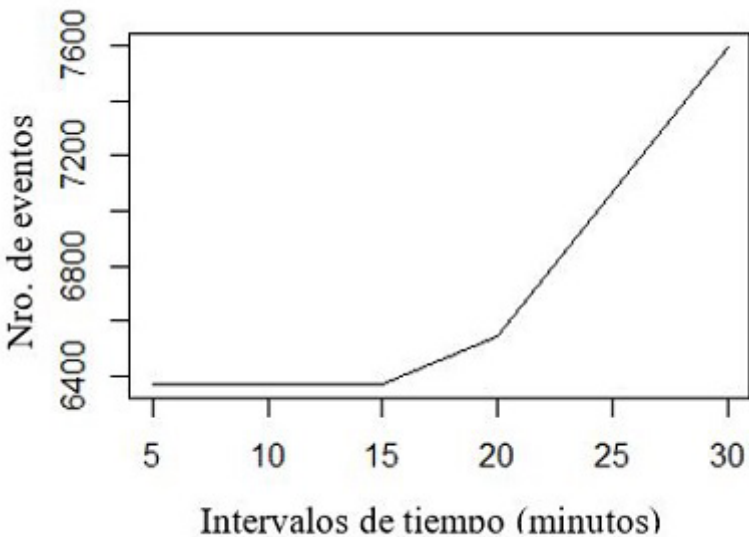


**Figura 7.** Comportamiento por año de las variables hora de inicio y equipo.

Fuente: Elaboración propia.

*Prueba comparativa entre el comportamiento de las variables en presencia de colisiones*

Una colisión está definida como el momento en el cual la hora final de una emergencia coincide con la hora de inicio siguiente en el mismo centro, el mismo día. La hora inicial que se registra es la hora que se despacha la unidad para atender la emergencia. La diferencia de tiempo que se toma para esta prueba parte del análisis de las colisiones que se presentan en diferentes intervalos de tiempo (ver figura 8). Se calculó la cantidad de eventos que se presentan si se toman distintos intervalos de tiempo, en este caso de cinco, diez, quince, veinte, veinticinco y treinta minutos, resultando quince minutos el valor tope que se obtiene antes de dispararse el comportamiento de la curva, por ende, se usó este intervalo para las pruebas correspondientes a colisiones.



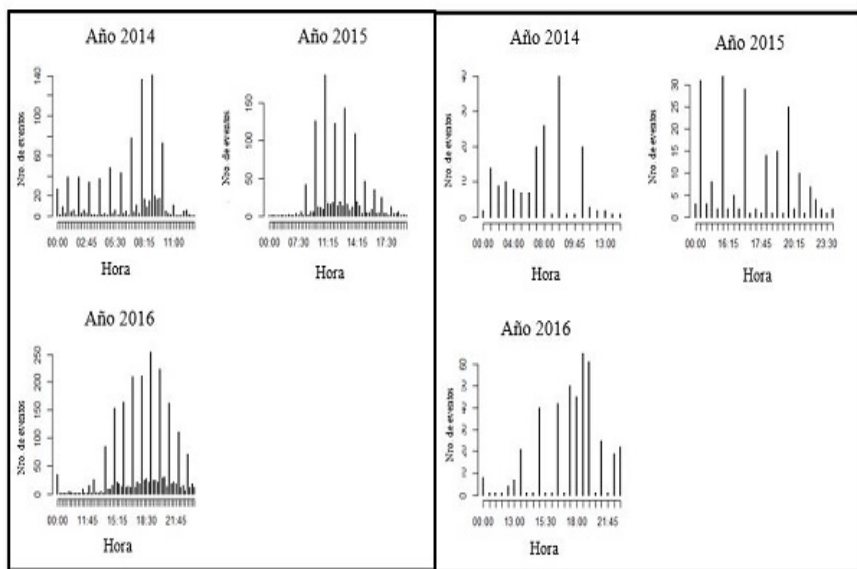
**Figura 8.** Curva de las colisiones en cada intervalo de tiempo.

Fuente: Elaboración propia.

Se compara el comportamiento por año de las actividades con respecto a cada variable con esta nueva prueba, en donde se obtiene cada variable a lo largo de cada año, pero con el añadido de que se tomaron solo las actividades que presentaron colisiones, es decir, aquellas cuya hora de inicio y hora final son menores al intervalo de quince minutos. En la figura 9 se presentan los resultados, a la izquierda se reflejan los resultados obtenidos en la prueba anterior, y a la derecha la nueva prueba considerando colisiones.

a) Hora de inicio sin colisión

b) Hora de inicio con colisión



**Figura 9.** Comparación de la variable hora de inicio para los tres años bajo estudio. (a) Sin colisión. (b) Con colisión. Fuente: Elaboración propia.

En el cuadro 3 se presenta una tabla comparativa considerando el número de la semana del año (tomando solo las tres semanas con mayor cantidad de emergencias presentadas en orden), la hora de inicio (colocada en orden de izquierda a derecha en el horario de 24 horas), los días, equipo y centros.



**Cuadro 3.** Comparación de variables sin y con colisiones.

Variable	Sin Colisión			Colisión		
	2014	2015	2016	2014	2015	2016
Nro. Semana	11 – 6 – 50	1 – 20 – 2	43– 10 – 11	50 – 19 – 6	27 – 1 – 18	24 – 48 – 25
Día	Jueves	Jueves	Sábado	Viernes	Jueves	Sábado
Hora Inicio	9 – 8 – 7	11– 13 – 10	19– 20 – 18	9 – 8 – 7	16 – 15 – 17	19 – 20 – 17
Equipo	SAR – Rojo	SAR – Rojo	SAR – Rojo	Sin unidad SAR	Sin unidad SAR	Sin unidad SAR
Centro	Panamericana Libertador Mocotíes	Panamericana Mocotíes Sur del Lago	Sur del Lago Panamericana Libertador	Panamericana Libertador Paramo	Sur del Lago Panamericana Mocotíes	Sur del Lago Libertador Panamericana

Fuente: Elaboración propia.

#### 4.4.3. Pruebas en torno a las colisiones

##### *Colisiones de horas en los diferentes centros*

A partir de la distribución de los centros a lo largo del Estado y, la manera en que se realiza el despliegue de unidades; es oportuno realizar un análisis de cada centro para las actividades no programadas. Esta prueba consistió en encontrar las colisiones tomando en cuenta solo aquellos centros que cubren las mismas zonas del Estado. Considerando los centros en donde se suscitan una mayor cantidad de emergencias se propone encontrar el comportamiento de las variables día y hora de inicio, dos variables de gran importancia al momento de encontrar información que permita a la organización tomar medidas preventivas para lograr un mejor servicio (ver figura 10).

##### *Horas en que se presentan las colisiones en cada centro durante el transcurso de la semana*

Se precisa en esta prueba las horas en donde se debe tener más equipo a disposición para poder cubrir todas las emergencias que se presentan en estos centros en particular. Cabe resaltar que no se puede globalizar

el comportamiento general de un centro (cada día presenta un comportamiento diferente) ni definir un patrón que permita generalizar el comportamiento del mismo (ver figura 11).

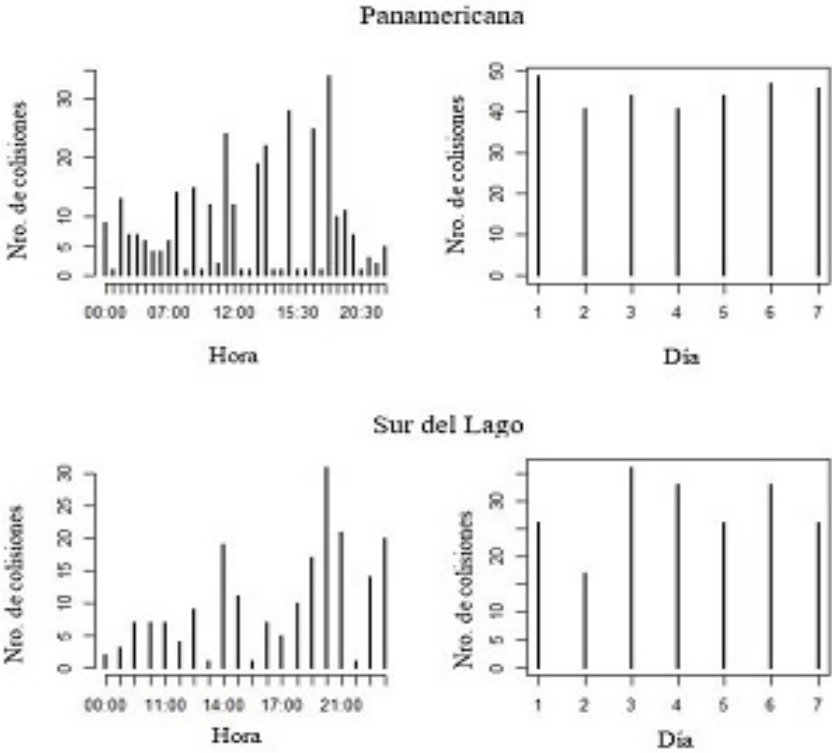


Figura 10. Centros Panamericana y Sur del Lago. Comportamiento de las colisiones para las variables hora y día. Fuente: Elaboración propia.

*Por actividad que presenta colisión se estudia para cada centro los equipos utilizados*

Las pruebas realizadas en torno a las colisiones permitieron observar que es importante buscar los equipos usados en cada centro para cada actividad y así poder determinar específicamente cuáles equipos deben estar disponibles con el objeto de mejorar el tiempo de atención en las emergencias. Considerando los centros que prestan la mayor cantidad

de atención a emergencias (Panamericana, Libertador, Sur del Lago y Mocotíes) se observa por emergencia los equipos que se utilizan cuando hay presencia de colisiones, mostrando los datos necesarios para determinar los equipos que son más usados al momento de atender cada actividad (ver figura 12). En el cuadro 4 se presentan los resultados obtenidos, se listan los dos equipos con mayor cantidad de emergencias atendidas.

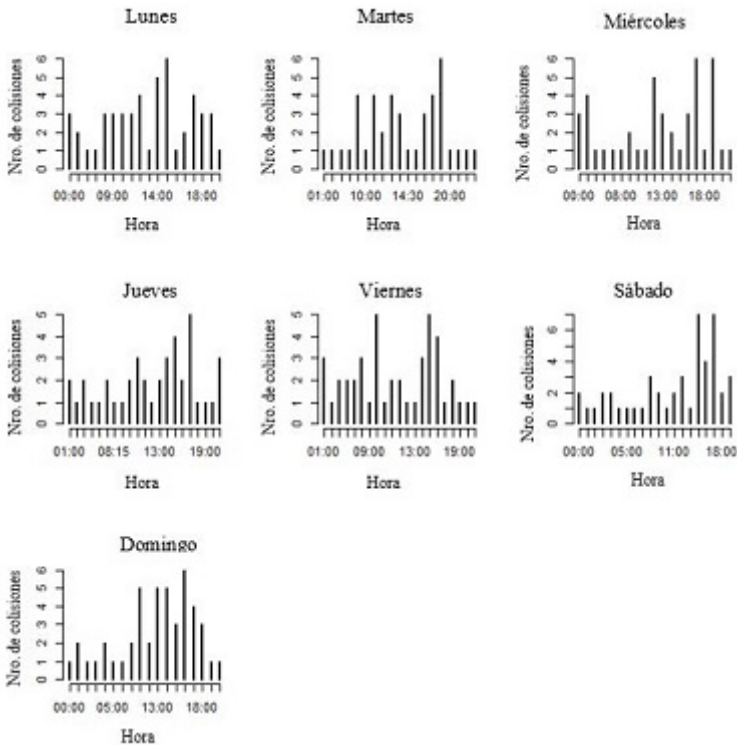
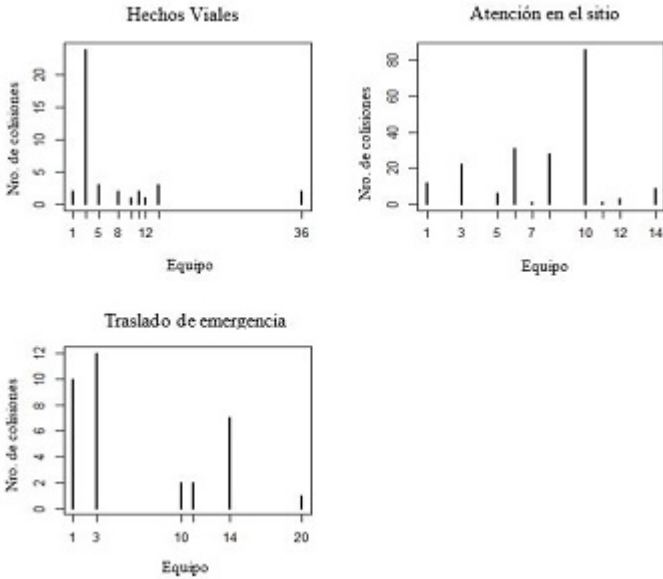


Figura 11. Centro Panamericana. Comportamiento de las horas durante cada día de la semana. Fuente: Elaboración propia.

Las pruebas relacionadas a colisiones pueden indicar si se necesitan más equipos, más personal, entre otros. Sin embargo, también es primordial saber el “cuándo”, al estudiar los datos por año.



**Figura 12.** Centro Libertador. Equipos utilizados para cada actividad no programada.  
Fuente: Elaboración propia.

**Cuadro 4.** Equipos usados en cada actividad no programada en los diferentes centros.

Centro	Actividad	Equipos
Sur del Lago	Hechos Viales	SAR – Rojo
	Atención en el sitio	Sin unidad – SAR
	Traslado de emergencia	Rojo – SAR
Panamericana	Hechos Viales	SAR – Rojo
	Atención en el sitio	Sin Unidad – SAR
	Traslado de emergencia	Rojo – SAR
Mocotíes	Hechos Viales	SAR – Rojo
	Atención en el sitio	Sin Unidad – SAR
	Traslado de emergencia	Rojo – SAR
Libertador	Hechos Viales	SAR – Halcón
	Atención en el sitio	Sin Unidad – Primeros Auxilios
	Traslado de emergencia	SAR – Rojo

Fuente: Elaboración propia.

#### 4.4.4. Prueba de datos tomando las semanas que presentan mayor cantidad de emergencias

Anteriormente se obtuvo el comportamiento global de las emergencias durante los tres años. Ahora, partiendo de la prueba inicial de las actividades no programadas, es prioritario brindar soporte durante aquellas semanas de alta demanda en el Estado Mérida en donde se presentan mayor cantidad de emergencias por año. En el cuadro 5 se presenta el comportamiento de las variables bajo estudio para dichas semanas.

**Cuadro 5.** Comportamiento de las variables durante las semanas que presentan mayor cantidad de emergencias.

Semana 2		Semana 11	
Variable	Valores	Variable	Valores
Día	Martes, jueves	Día	Lunes, viernes
Hora de Inicio	7:00 p.m.	Hora de Inicio	8:15 p.m.
Equipo	SAR, Rojo	Equipo	SAR, Rojo
Centro	Paramo, Mocoties, Panamericana	Centro	Sur del Lago, Panamericana
Semana 37 y 38		Semana 43	
Variable	Valores	Variable	Valores
Día	Viernes, lunes	Día	Miércoles
Hora de Inicio	8:15 pm, 3:15 pm	Hora de Inicio	1:30 pm, 9:00 am
Equipo	SAR - Rojo	Equipo	SAR, Rojo
Centro	Panamericana, Sur del Lago	Centro	Sur del Lago, Panamericana

Fuente: Elaboración propia.

#### 4.4.5. Prueba de datos durante los períodos vacacionales

Mérida por ser un Estado turístico presenta gran afluencia de visitantes durante los períodos vacacionales (semana santa, agosto y la temporada navideña). Particularmente, la ciudad de Mérida recibe visitantes para las Ferias del Sol. Por tal motivo, estas cuatro temporadas y la variable Municipio se tomaron en cuenta para realizar dicha prueba. Esta

última variable se excluyó de la temporada de ferias debido a que estas se desarrollan solo en el Municipio Libertador (ver cuadros 6 y 7, respectivamente).

**Cuadro 6.** Ferias del Sol.

Ferias del Sol	
Variable	Valores
Día	Lunes y jueves
Equipo	SAR
Hora de Inicio	3:30 p.m.

Fuente: Elaboración propia.

**Cuadro 7.** Épocas vacacionales.

	Época navideña	Semana Santa	Temporada de agosto
Variable	Valores	Valores	Valores
Día	Jueves	Lunes	Sábado
Hora Inicio	5:00, 8:00 y 9:00 pm	4:00 y 6:00 pm	7:00 y 6:00 pm
Equipo	SAR, Rojo	Rojo, SAR	SAR, Rojo
Centro	Panamericana	Sur del Lago	Panamericana
	Libertador	Libertador	Sur del Lago
Municipio	Tovar	Tulio Febres Cordero	Tovar
	Caracciolo Parra	Libertador	Caracciolo Parra

Fuente: Elaboración propia.

## 5. Resultados

Los resultados se presentan en cinco secciones, de acuerdo al orden contemplado en la metodología.

### 5.1. Análisis de las pruebas iniciales

Las pruebas iniciales de frecuencia para cada variable orientaron la atención hacia la variable horas, observándose un comportamiento de

“horario de oficina”, es decir, picos altos a las ocho, nueve y diez de la mañana, disminución en horas de almuerzo y luego, otro pico al comienzo de la tarde. Estos resultados llevaron a realizar una segunda prueba, la cual se enfocó en la búsqueda por semana de algún día en particular que presentara un comportamiento diferente al de “oficina”, incluso los días viernes y sábado se comportan de la misma forma (ver figura 13).

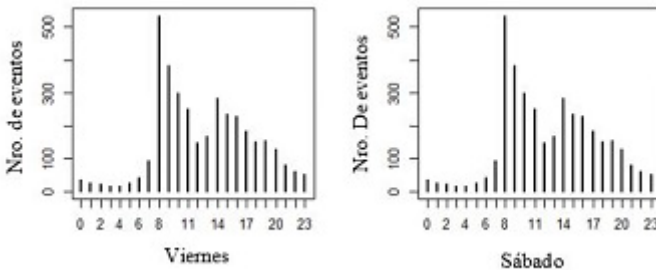


Figura 13. Comportamiento días viernes y sábado.

Fuente: Elaboración propia.

En tal sentido, el resultado poco satisfactorio en las pruebas iniciales conduce a realizar una prueba de las horas en que se presentan los eventos, pero tomando en cuenta las actividades como eje central, encontrándose que las actividades que requieren planeación previa son aquellas que se dan en ese llamado “horario de oficina”, y por ser estas las actividades que más desempeña la organización estaban afectando la prueba general.

## 5.2. Análisis de las pruebas en torno a las actividades no programadas

En la prueba general realizada a las actividades no programadas resaltan las variables centro y hora de inicio. Esta última variable llama la atención, debido a que en la prueba anterior se presentaba en “horario de oficina”, pero ahora aislando las emergencias, se presentan en su mayoría en un horario tarde-noche. Con la nueva información, se decidió realizar una

prueba por año de las emergencias para ver si reflejaban cambios en su comportamiento durante los tres años bajo estudio. Dicha prueba arrojó que las variables no se comportaban igual para ningún año a excepción de la variable equipo, donde SAR y Rojo se presentan siempre como los dos equipos más usados en la atención de emergencias (ver figura 14).

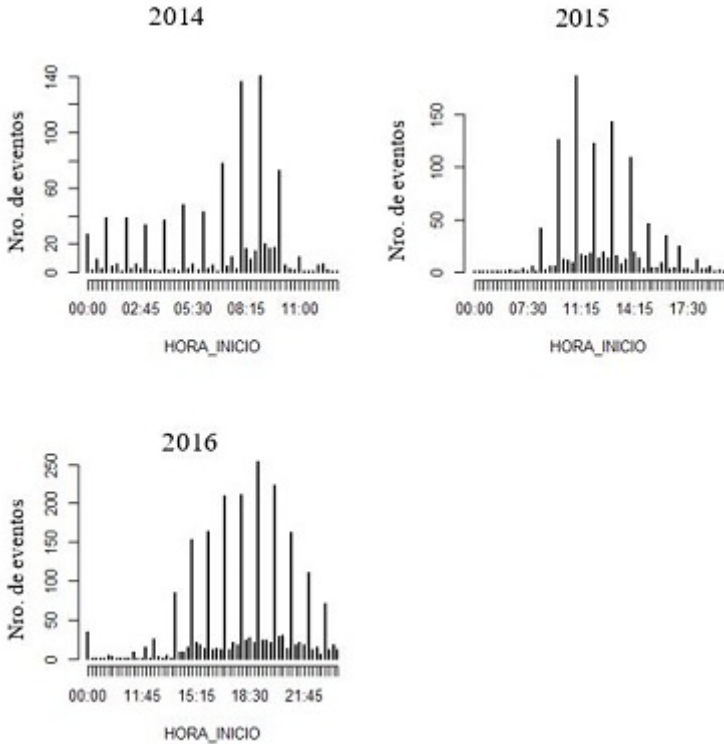


Figura 14. Horas de las emergencias presentadas por año.

Fuente: Elaboración propia.

Los resultados anteriores permitieron obtener un esquema de comportamiento general de cada variable (ver figura 15).





Figura 15. Esquema de comportamiento de las emergencias.

Fuente: Elaboración propia.

### 5.3. Análisis de las pruebas en torno a las colisiones

La prueba de las colisiones en los diferentes centros para las variables hora y día arrojó que dichas colisiones, a pesar de ser relativamente pocas por centros, existen y son casos en las que el retardo en la atención de emergencia puede arriesgar la vida de las personas. Por otra parte, la prueba conjunta de las variables hora – día (considerando aquellos centros que atienden un mayor número de emergencias) permitió

definir bloques de horas para cada municipio donde se presentan mayores colisiones.

En el cuadro 8 se presentan los puntos críticos por día de la semana para el centro Libertador, observándose lo siguiente:

- Colisiones en el bloque del llamado “salida del trabajo” o “regreso a casa” entre las seis y siete de la noche. Además, al acercarse el fin de semana las colisiones pasan a ser una hora más tarde, entre las siete y ocho de la noche.
- Entre semana se observó un patrón indicando existencia de colisiones a las dos de la tarde, el cual, al relacionarse con el estilo de vida del municipio al que atiende concuerda con la hora después de almuerzo en que la ciudadanía regresa a trabajar para el turno laboral vespertino.

**Cuadro 8.** Colisiones para el centro Libertador.

Hora (24 horas) / Día	Centro Libertador												
	8	9	10	11	12	13	14	15	16	17	18	19	20
Lunes													
Martes													
Miércoles													
Jueves													
Viernes													
Sábado													
Domingo													

Fuente: Elaboración propia.

En el cuadro 9 se presentan los puntos críticos para el centro Panamericana, uno de los centros encargado de toda la demanda que se presenta en la carretera que conecta a la ciudad de Mérida con la de El Vigía, ubicado en Chiguará. En este caso, se observó la presencia de tres bloques en diferentes horas durante el día.

- Bloque diez de la mañana.

- Bloque mediodía- tarde, que va desde la una hasta las dos de la tarde en su mayoría, con excepción del viernes y sábado donde las colisiones se presentan a las tres de la tarde.
- Un último bloque en la tarde - noche, que va desde las cinco hasta las siete de la noche.

**Cuadro 9.** Colisiones para el centro Panamericana.

Hora (24 horas) / Día	Centro Panamericana													
	8	9	10	11	12	13	14	15	16	17	18	19	20	
Lunes														
Martes														
Miércoles														
Jueves														
Viernes														
Sábado														
Domingo														

Fuente: Elaboración propia.

Finalmente, en el cuadro 10 se reflejan los resultados para el centro Mocotíes, el cual está encargado de atender las poblaciones de Tovar y Zea, observándose la presencia de dos bloques.

- Primer Bloque: a las once de la mañana con la excepción del día jueves.
- Segundo Bloque: en horas de la tarde, en su mayoría a las cinco.

En cuanto a los equipos se observó que las SAR (camionetas pick-ups) y Rojo (ambulancias) son los más usados para todos los centros, siendo necesario disponer de más equipos ya que los resultados indican que la emergencia "Atención en el sitio" particularmente se atiende sin unidad, es decir, no envían ningún equipo, solo oficiales, esto requiere que otro organismo haga el traslado en caso de ser necesario.

**Cuadro 10.** Colisiones para el centro Mocotíes.

Hora (24 horas) / Día	Centro Mocotíes												
	8	9	10	11	12	13	14	15	16	17	18	19	20
Lunes													
Martes													
Miércoles													
Jueves													
Viernes													
Sábado													
Domingo													

Fuente: Elaboración propia.

#### 5.4. Análisis de las semanas con mayor cantidad de emergencias

Los resultados obtenidos se presentan en el cuadro 11, en donde se clasifican las horas en tres bloques y se exponen los centros que más servicios prestaron durante esas semanas en particular, observándose lo siguiente:

- El comportamiento de las variables para las diferentes semanas varía con excepción de la variable equipo, que siempre para el caso de atención a emergencias es SAR y Rojo.
- Para las semanas con mayor demanda, las emergencias ocurren en un horario tarde-noche. Al enlazar ese dato con la variable centro se observó que Panamericana hace presencia en todas las semanas, aportando una posible relación entre esas horas y el centro.

La información resumida en dicho cuadro puede ser utilizada como herramienta para tomar decisiones en cuanto a la distribución de recursos humanos y equipos para estas semanas. Por ejemplo, la semana dos, tiene como centros Panamericana, Mocotíes y Páramo, lo cual se debe a que Mérida por ser una ciudad turística atiende a muchos visitantes durante la época navideña, y es alrededor de esta semana (primeros días de enero) que regresan a sus hogares. Justamente, los

centros con mayor demanda son aquellos que cubren la ciudad de Mérida y las zonas por donde existen vías usadas para entrar y salir de la ciudad, es decir, Páramo y Panamericana. Finalmente, se observó que Panamericana es un centro de atención crítico ya que durante todas las semanas bajo estudio está presente con altos números de atención a emergencias.

**Cuadro 11.** Comportamiento de las variables para las semanas con más eventos.

	Comportamiento de las semanas con más emergencias													
	Hora (24 horas)			Día					EQ	Centro				
	8-sept	13-15	19-20	L	M	MI	J	V	S	SAR Rojo	Pana.	Mocot.	SurLa.	Páramo
Semana 2														
Semana 11														
Semana 37 y 38														
Semana 10														
Semana 43														

Fuente: Elaboración propia.

### 5.5. Análisis de las emergencias en temporadas vacacionales

Para la temporada de ferias, los días con las emergencias son los jueves y lunes de carnaval, la hora en la que más se presentan es alrededor de las 3:30pm. En cuanto a las temporadas vacacionales se observaron dos datos resaltantes: el día sábado está presente en todas las temporadas; así como las horas tarde-noche en las que se presentan las emergencias coinciden para estas semanas (ver cuadro 12).

**Cuadro 12.** Comportamiento de variables día y hora para temporadas vacacionales.

	Día						Hora (Formato de 24 horas)					
	L	M	MI	J	V	S	D	4-jun	8-oct	14-16	17-19	20-22
Navidad												
Semana Santa												
Agosto												

Fuente: Elaboración propia.

En general, se observó que la temporada navideña está marcada por presentar emergencias en la tarde-noche, en el centro que sirve a la ciudad de Mérida y sus alrededores, y como característica particular las emergencias ocurren en su mayoría los fines de semanas. En Semana Santa las emergencias se presentan entre semana con la excepción del sábado, y las emergencias comienzan a presentarse en un horario más temprano que el navideño, pero se extiende tanto como el anterior. En el caso de los centros se comporta igual que la temporada navideña con el agregado del centro Sur del Lago.

Por su parte, la temporada vacacional se caracteriza por ser la más larga entre todos los períodos vacacionales que se presentan a lo largo del año, en este caso los centros varían un poco, además ya no está presente el centro Libertador, lo que lleva a suponer que las personas salen de la ciudad de Mérida a vacacionar a otros lugares, o la afluencia de turistas se dirige a otros lugares del Estado. Las horas en las que se presentan emergencias se comportan de manera muy parecida a la temporada navideña para la tarde-noche, pero agosto presenta la particularidad de que es el único período vacacional que presenta una alta cantidad de emergencias en horas de la mañana.

## 6. Conclusiones

Se implementaron algoritmos no supervisados o de descubrimiento del conocimiento con el objeto de encontrar patrones o tendencias que llevaran a la extracción valiosa de información, escribiendo algoritmos de listas,

búsqueda y filtrado de datos por medio de la creación de diferentes scripts para la realización de las distintas pruebas que se llevaron a cabo. Los resultados de las pruebas iniciales condujeron el trabajo a enfocarse en las actividades no programadas, es decir, las emergencias que se presentan en el Estado Mérida.

Las pruebas en torno a las emergencias se enfocaron en aquellas horas donde es probable que la emergencia no haya sido atendida inmediatamente, a estos casos se les llamó colisiones. Por otra parte, se llevó a cabo una prueba de las emergencias centrada en encontrar comportamientos de las variables durante las colisiones, encontrándose las horas en las cuales se suscitan colisiones en cantidades relevantes, de modo que la organización ahora conozca y pueda estar alerta a estas determinadas horas. Asimismo, se encontró que el equipo SAR y Rojo son los más usados en la atención de emergencias; sin embargo, en algunos casos, durante las colisiones dichas emergencias eran atendidas sin equipo alguno y solo despachaban oficiales para atenderlas. En tal sentido, a partir de la cantidad de colisiones que se presentan para un mismo equipo se concluyó que es necesaria una gestión que disponga de mejor manera los equipos y el personal.

El estudio en torno a las semanas con mayor cantidad de emergencias y, de las semanas que conforman las temporadas vacacionales durante el año, arrojó que las emergencias ocurren en su mayoría en un horario de tarde-noche a excepción de algunos casos puntuales, como las vacaciones de agosto en donde se presentan también en horas de la mañana. Por otra parte, al considerar algunos valores específicos para la variable semana se observó que el día sábado presenta la mayor cantidad de emergencias seguido del día lunes, y claramente se obtuvo un intervalo de hora con mayor demanda entre las cinco y diez de la noche para las temporadas vacacionales. Adicionalmente, se observó que el centro que más servicio presta es Panamericana, esto brinda información a la organización que pueden usar al momento de crear algún plan de gestión para temporadas vacacionales específicas.

INPRADEM como organización que trabaja en pro de la comunidad merideña, se beneficia de este proyecto al conocer los puntos críticos

en los cuales tienen mayor demanda sus servicios a emergencias, esto puede permitir crear un plan de gestión que vaya mejor orientado y sea más preciso, ya que ahora cuenta con información que indica el comportamiento de las semanas durante el año en que se necesita estar más alerta, así como también los días y las horas en las que más se presentan emergencias durante la semana, sin olvidar el centro que mayor atención presta a los hechos ocurridos en el Estado. En tal sentido, se recomienda la creación de una planilla diseñada por un experto, que permita una extracción de información de mayor calidad para ser usada en análisis futuros. Asimismo, otorgar a cada funcionario un código específico que permita realizar pruebas desde el punto de vista de recursos humanos. Por otra parte, asignar código a cada unidad que forma parte del código general de la variable equipo, por ejemplo: una moto asignada al centro Libertador, con características como color o placa, no debe simplemente estar generalizada bajo el código Halcón, asignado para clasificar la flota de motos con el que cuenta la organización, debe poseer un código específico que permita conocer que esa moto en particular fue usada; esto sería ideal para obtener con detalles una prueba de los equipos utilizados en la atención de emergencias, permitiendo mejorar la gestión al saber con exactitud en qué momento se puede prescindir de un equipo en particular en un centro y asignarlo a otro.

Es necesario recordar que un proyecto de minería de datos centrado en ayudar en la toma de decisiones de una organización, debe considerar el nivel de logro de objetivos como satisfactorios si los resultados obtenidos ayudan a la organización a cumplir sus objetivos o metas establecidas. Tres factores cualitativos que determinan que tan satisfactorio fue el estudio son la precisión, la utilidad y el beneficio obtenido del conocimiento adquirido. Este proyecto basado en la aplicación de minería de datos, que tiene como objetivo realizar la descripción de datos, específicamente caracterizar, colabora como soporte a la toma de decisiones de esta organización, al brindar información que permita optimizar la distribución de equipo y personal, disminuyendo posiblemente el tiempo de espera de una emergencia y mejorando la calidad de atención de estas, al tener a disposición equipos necesarios para el servicio que brinda en presencia de emergencias suscitadas a lo



largo del Estado Mérida. El nivel de precisión es bastante satisfactorio considerando que solo se cuentan con tres años para la realización de este estudio, aun así, ahora INPRADEM cuenta con una nueva forma de ver sus datos, prometiendo beneficios a la solución de una gran variedad de problemas como: planeación económica, distribución de trabajo y equipo, análisis de servicio y prevención en épocas de demanda alta.

Siguiendo la metodología CRISP-DM, su última fase, despliegue de resultados, se refiere a todo lo relacionado a la aplicación del conocimiento adquirido en el estudio, así como el reporte final, pero en el caso particular de este estudio, la información encontrada no puede ser implementada inmediatamente, y la evaluación de su nivel de utilidad debe provenir desde la organización, ya que el desarrollador de este trabajo no forma parte de la institución y la aplicabilidad de esta información depende de las personas que hacen vida en la organización a la que le será entregada una copia de este proyecto.

## 7. Notas

1. La minería de datos ha aumentado su presencia en los últimos años por ofrecer en un sinfín de campos una gran variedad de aplicaciones. Una de las ramas que ha sacado más provecho de la minería de datos es sin duda la rama empresarial, las organizaciones siempre en la búsqueda de incrementar beneficios y ganancias desean controlar las operaciones y servicios que ofrecen para mejorar de manera continua. Las organizaciones públicas, las cuales son establecidas por el gobierno para proporcionar servicios públicos, comparadas con las empresas privadas, buscan servir y ser de utilidad general a la colectividad a la que pertenecen. En la búsqueda de ese objetivo, una organización pública como INPRADEM, con una trayectoria de compromiso con la población, se beneficiará de la minería de datos mediante la obtención de información para la elaboración de futuros planes, que contribuirán en mejoras en el proceso de gestión y toma de decisiones.

2. Todos los demás resultados junto con los Scripts creados, así como la base de datos con la cual se trabajó, se encuentran en una carpeta pública de Google Drive bajo la siguiente dirección: [https://drive.google.com/open?id=1WDJrLKAiTcNVcUcfMs\\_KYMTld118xR1W](https://drive.google.com/open?id=1WDJrLKAiTcNVcUcfMs_KYMTld118xR1W).

## 8. Referencias

- Aránguiz, Alejandro (2012). *Análisis de accidentes de tránsito en zonas urbanas y rurales usando minería de datos difusa*. Tesis de Maestría. Chile: Pontificia Universidad Católica de Valparaíso, 2012, 104 pp.
- Chapman, Pete; Clinton, Julian; Kerber, Randy; Khabaza, Thomas; Reinartz, Thomas; Shearer, Colin y Wirth, Rudiger (2000). *CRISP-DM 1.0 Step-by-step Data Mining Guide*. Consultado: 02/04/2017. Disponible (online): <https://the-modeling-agency.com/crisp-dm>.
- García, René; Delgado, Domingo; Díaz, Eduardo y García, René (2012). "Caracterización de la accidentalidad vehicular y análisis de las causas en la provincia de Villa Clara, Cuba". *Dyna*, 79, 175 (octubre, 2012), pp. 191-200. Consultado: 12/04/2017. Disponible (online): <http://revistas.unal.edu.co/index.php/dyna/article/view/20038/43525>.
- Emergencia 171 Mérida (2018-2019). *Protección civil*. Consultado: 22/03/2017. Disponible (online): <http://171.merida.gob.ve/proteccion-civil>.
- Govindarajan, Muthukumarasamy y Chandrasekaran, Ramanathan (2011). "Intrusion detection using neural based hybrid classification methods". *Computer networks*, 55, 8 (junio, 2011), pp. 1662-1671. DOI: <https://doi.org/10.1016/j.comnet.2010.12.008>.
- Instituto Nacional de Estadística (INE) (2013). *Cifras del parque automotor del estado Mérida*. Consultado: 03/04/2017. Disponible (online): <http://www.ine.gov.ve/>.
- Larose, Daniel y Larose, Chantal (2014). *Discovering knowledge in data: an introduction to data mining*. New Jersey: John Wiley y Sons, pp. 336.

- León, Elizabeth (2017). *Módulo minería de datos*. Consultado: 17/05/2017. Disponible (online): [http://disi.unal.edu.co/~eleonguz/cursos/md/presentaciones/Sesion5\\_Metodologias.pdf](http://disi.unal.edu.co/~eleonguz/cursos/md/presentaciones/Sesion5_Metodologias.pdf).
- Hassinger, Mark (2015). *Aplicación de técnicas de minería de datos en accidentes de tráfico*. Tesis de maestría. Valencia: Universidad Politécnica de Valencia, 2015, 67 pp. Consultado: 05/04/2017. Disponible (online): <https://riunet.upv.es/handle/10251/65082>.
- Salinas, Marcos y Vele, Luis (2014). *Estudio científico de la accidentalidad de tránsito en El Cantón Cuenca*. Tesis de pregrado. Ecuador: Universidad Politécnica Salesiana Sede Cuenca, 2014, 117 pp. Consultado: 22/03/2017. Disponible (online): <https://dspace.ups.edu.ec/handle/123456789/7209>.
- Timarán, Silvio; Hernández, Isabel; Caicedo, Segundo; Hidalgo, Arsenio y Alvarado, Juan (2016). *Descubrimiento de patrones de desempeño académico con árboles de decisión en las competencias genéricas de la formación profesional*. Bogotá: Ediciones Universidad Cooperativa de Colombia, 193 pp. DOI: 10.16925/9789587600490.
- Villarino, Guillermo (2015). *Metodología de minerías de datos para el estudio de tablas de siniestralidad vial*. Tesis de maestría. Madrid, España: Universidad Complutense, 91 pp.
- Zavala, Francisco (2014). *Buscador de artículos científicos aplicando minería de texto*. Tesis de maestría. México: Instituto Tecnológico de la Paz, 2014, 95 pp. Consultado: 24/04/2017. Disponible (online): <https://docplayer.es/3887260-Buscador-de-articulos-cientificos-aplicando-mineria-de-datos-tesis.html>.