



PROYECTO DE GRADO

CARACTERIZACIÓN DE LOS REGISTROS DE EMERGENCIAS
ATENDIDAS POR INPRADEM UTILIZANDO MINERÍA DE DATOS

www.bdigital.ula.ve

Por

Br. Yamilet Casanova

Tutores: Prof. Lucileima Rosales

Prof. Francisco Hidrobo

Mayo 2018

©2018 Universidad de Los Andes Mérida, Venezuela

C.C. Reconocimiento

CARACTERIZACIÓN DE LOS REGISTROS DE EMERGENCIAS ATENDIDAS POR INPRADEM UTILIZANDO MINERÍA DE DATOS.

Br. Yamilet Casanova.

Proyecto de grado – Investigación de operaciones, 74 páginas.

Resumen: Durante los últimos años el crecimiento poblacional del Estado Mérida y su carácter de Estado turístico, trae consigo que se requiera un constante apoyo y servicio de Instituciones que estén alerta ante posibles contratiempos. En vista de estas necesidades la Gobernación creó Protección Civil y Administración de Desastres del Estado Mérida (INPRADEM), organismo de carácter público con sede principal en la Av. Los Próceres, Sector Santa Bárbara Oeste de la ciudad de Mérida. En aras de mejorar los procesos internos de la Institución, este proyecto, siguiendo la metodología CRISP–DM, obtiene información que puede ser de utilidad a la Institución como soporte en la toma de decisiones. Los datos se analizaron por medio de un conjunto de pruebas, divididas en cinco secciones: la primera orientada al análisis general de los datos, la segunda a las emergencias suscitadas en el Estado, la tercera centrada en colisiones de horas, la cuarta pruebas entorno a las semanas en donde se reportan mayor cantidad de emergencias, y por último, las relacionadas con los períodos de las temporadas vacacionales. Los resultados aportan información relevante: se encontraron las horas en las cuales se suscitan colisiones. Así mismo, se encontró que el equipo SAR y Rojo son los más usados en la atención de emergencias, pero durante las colisiones se atienden emergencias sin equipo alguno y solo son despachados oficiales. Las emergencias ocurren en su mayoría en un horario tarde-noche a excepción de las vacaciones de agosto en donde se presentan también en horas de la mañana. En temporadas vacacionales el día sábado presenta la mayor cantidad de emergencias. En cuanto a las horas con mayor demanda se encuentra un intervalo entre las cinco y diez de la noche. Para los centros se encontró que Panamericana es el más servicio presta. 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 clave: Organización, Minería de Datos, Caracterización.

INDICE GENERAL

ÍNDICE DE FIGURAS	5
ÍNDICE DE TABLAS	7
CAPÍTULO I. EL PROBLEMA	8
1.1. Introducción	8
1.2. Planteamiento del Problema.....	9
1.3. Objetivos	10
1.3.1 Objetivo General.....	10
1.3.2 Objetivos Específicos	10
1.4. Antecedentes	10
CAPÍTULO II. MARCO TEÓRICO.....	13
1.1. Introducción	13
1.2. Organización	14
1.2.1 Organización pública.....	14
1.2.2 INPRADEM	14
2.3. Minería de Datos	16
2.3.1. Definición	16
2.3.2. Áreas de aplicación.....	16
2.3.3. Objetivos de la Minería de Datos	17
2.3.4. Metodologías	17
2.4. Metodología CRISP-DM.....	19
2.5. Minería de datos en las organizaciones.....	21
2.5.1. Minería de datos en los procesos de toma de decisiones.....	21
3.1. Introducción	23

3.2.	Fases 1 y 2 de la Metodología CRISP-DM	23
3.2.1.	Compresión y análisis del problema.....	23
3.2.2.	Entendimiento de datos	23
3.3.	Fase 3: Preparación de los datos	24
3.4.	Fase 4: Pruebas iniciales	30
3.4.1.	Gráfico de frecuencia para cada variable	30
3.4.2.	Por día las horas en que se presentan los eventos	35
3.4.3.	De las actividades que más se realizan en la organización se estudia el comportamiento de las horas en que se presta el servicio	36
3.5.	Pruebas entorno a las actividades no programadas	38
3.5.1.	Comportamiento general de actividades con respecto a cada variable	38
3.5.2.	Comportamiento por año de las actividades con respecto a cada variable	39
3.5.3.	Prueba comparativa entre el comportamiento de las variables en presencia de colisiones durante los años bajo estudio.....	40
3.6.	Pruebas entorno a las colisiones.....	43
3.6.1.	Colisiones de horas en los diferentes centros	43
3.6.2.	Horas en que se presentan las colisiones en cada centro durante el transcurso de la semana.....	44
3.6.3.	Por actividad que presenta colisión se estudia para cada centro los equipos utilizados.....	46
3.7.	Prueba de datos tomando las semanas que presentan mayor cantidad de emergencias.....	48
3.8.	Prueba de datos durante los períodos vacacionales.....	50
CAPÍTULO 4. ANÁLISIS DE RESULTADOS		54
4.1.	Análisis de pruebas iniciales	54

4.2.	Análisis de pruebas entorno a las actividades no programadas	55
4.3.	Análisis de los pruebas entorno a las colisiones	57
4.4.	Análisis de las semanas con mayor cantidad de emergencias.....	61
4.5.	Análisis de las emergencias en temporadas vacacionales	63
CAPÍTULO V. CONCLUSIONES Y RECOMENDACIONES		66
5.1.	Conclusiones	66
5.2.	Recomendaciones.....	69
REFERENCIAS		71

www.bdigital.ula.ve

ÍNDICE DE FIGURAS

Figura 1. Fases de la metodología SEMMA	17
Figura 2. Fases de la metodología KDD	18
Figura 3. Visión general CRISP-DM	18
Figura 4. Fases de la Metodología CRISP-DM.....	19
Figura 5. Esquema de clasificación de las variables	29
Figura 6. Frecuencias para las variables Semana, Día, Municipio y Actividad	34
Figura 7. Hora en que presentan los eventos para los días lunes, martes, miércoles y jueves	35
Figura 8. Horas en que se presta el servicio para diferentes actividades.....	37
Figura 9. Comportamiento por año de las variables Hora de Inicio y Equipo de las actividades no programadas.....	39
Figura 10. Curva de las colisiones en cada intervalo de tiempo.....	40
Figura 11. Comparación de la variable hora de inicio para los tres años bajo prueba con o sin prueba de colisión	41
Figura 12. Para los centros Panamericana y Sur del Lago, comportamiento de las colisiones para las variables Hora y Día.....	43
Figura 13. Comportamiento de las horas durante cada día de la semana para el centro Panamericana.....	45
Figura 14. Equipos utilizados para cada actividad no programada en el centro Libertador.	46
Figura 15. Comportamiento de las variables para la semana número dos del año.....	48
Figura 16. Comportamiento de las variables para la semana número once del año.....	49
Figura 17. Comportamiento de las variables para la época navideño y la temporada vacacional de agosto.....	52
Figura 18. Comportamiento de los días viernes y sábado para la prueba 3.4.2	54
Figura 19. Horas de las emergencias presentadas para cada año	56
Figura 20. Esquema de comportamiento de las emergencias.....	56
Figura 21. Para los centros Libertador y Pueblos del Sur, comportamiento de las colisiones para las variables Hora y Día del prueba 3.6.1	57

Figura 22. Comportamiento de las horas durante cada día de la semana para el centro Libertador del prueba 3.6.2.....	57
Figura 23. Resumen de resultados para la temporada navideña.....	64
Figura 24. Resumen de resultados para la temporada de Semana Santa.....	65
Figura 25. Resumen de resultados para la temporada vacacional de agosto.....	65

www.bdigital.ula.ve

ÍNDICE DE TABLAS

Tabla 1. Variables con su descripción	24
Tabla 2. Bases de datos inicial.....	25
Tabla 3. Datos después del primer filtrado.....	27
Tabla 4. Ejemplos de variables llevadas a su código numeral	27
Tabla 5. Datos después de asignarle su código numeral	28
Tabla 6. Variable semana	30
Tabla 7. Variable día	30
Tabla 8. Variable Hora de Inicio	31
Tabla 9. Variable Hora Final	31
Tabla 10. Variable Nro. de Funcionarios	31
Tabla 11. Variable Actividad.....	32
Tabla 12. Variable Departamento.....	32
Tabla 13. Variable Equipo.....	33
Tabla 14. Variable Municipio.....	33
Tabla 15. Comportamiento de las actividades no programadas con respecto a cada variable	38
Tabla 16. Tabla comparativa de las variables con o sin prueba de colisiones	42
Tabla 17. Equipos usados en cada actividad no programada en los diferentes centros	47
Tabla 18. Comportamiento de las variables durante las semanas que presentan mayor cantidad de emergencias.....	50
Tabla 19. Resultado de las variables para la prueba de las Ferias del Sol.....	51
Tabla 20. Valores de las variables obtenidos en el prueba de las épocas vacacionales	51
Tabla 21. Tabla de colisiones para el centro Libertador.....	59
Tabla 22. Tabla de colisiones para el centro Panamericana	60
Tabla 23. Tabla de colisiones para el centro Mocotíes.....	60
Tabla 24. Comportamiento de las variables para las semanas con más eventos	62
Tabla 25. Comportamiento de variables día y hora para las temporadas vacacionales.....	63
Tabla 26. Comportamiento de la variable Centro para las temporadas vacacionales	64

CAPÍTULO I. EL PROBLEMA

1.1. 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 contribuye 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 (ej. 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 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), organismo de carácter público, social y humanitario, creado ante la eminente situación de riesgo natural y elevada incidencia de accidentes antrópicos que caracteriza el Estado. Tiene su ubicación geográfica en el Estado Mérida, sede principal Av. Los Próceres, Sector Santa Bárbara Oeste.

Este organismo ejecuta 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.

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.

En tal sentido, dichos datos pueden ser analizados para aportar información útil que contribuya a la mejora de la gestión de la Organización mediante la minería de datos, la cual es una técnica que se encarga de la extracción no trivial de información implícita, previamente desconocida y potencialmente útil a partir de datos con el fin de descubrir patrones con significado.

En este trabajo se pretende caracterizar los datos registrados de las emergencias atendidas por INPRADEM, mediante técnicas de minería de datos, para dar soporte a la toma de decisiones en los procesos de gestión de la Organización.

1.2. Planteamiento del Problema

El Instituto de Protección Civil y Administración de Desastres del Estado Mérida (INPRADEM), cuyo objetivo y compromiso con la ciudadanía acentúa la disposición y ganas de mejorar su servicio de atención en las emergencias, mediante la recolección de datos por medio de una planilla que se almacena de forma mensual y por sectores, en donde hay una filial del organismo, indicando los detalles de la unidad despachada y de la emergencia atendida, la planilla posee datos bien organizados y completos de cada emergencia atendida por la Organización.

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.

1.3. Objetivos

1.3.1 Objetivo General

Caracterizar los datos registrados de las emergencias atendidas por INPRADEM, mediante técnicas de minería de datos, para dar soporte a la toma de decisiones en los procesos de gestión de la Organización.

1.3.2 Objetivos Específicos

- Estudiar los formatos de recolección de información actuales y sus esquemas de almacenamiento.
- Modificar los esquemas de representación de los datos para que sean utilizables con las herramientas de minería de datos a usar.
- Definir las estrategias de caracterización de los datos y su representación.
- Implementar las técnicas adecuadas que permitan la caracterización.
- Describir los ámbitos de aplicabilidad de la información obtenida.

1.4. Antecedentes

Aránguiz (2012), realizó un prueba basado en la aplicación de técnicas de minería de datos de segmentación, asociación y clasificación, orientadas en la identificación de los patrones y variables más influyentes en la definición de un accidente de tránsito. Por otro lado, efectuó un análisis difuso del problema, a través de la creación de atributos difusos, para brindar mayor información sobre los accidentes, acompañados de algoritmos difusos, particularmente de clasificación. Para dicho 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.

Hassinger (2015), utilizó árboles de decisión para analizar los accidentes de tráfico, 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 los accidentes de tráfico.

Salinas y Vele (2014), realizaron una predicción de la tendencia de la cantidad de accidentes de tránsito registrados desde agosto 2014 hasta el año 2016. Posteriormente, mediante indicadores de accidentalidad de tránsito, abordaron un análisis descriptivo. Asimismo, 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, se formularon leyes, reformas y campañas enfocadas en la disminución de accidentes de tránsito en el cantón Cuenca.

García et al. (2012), realizaron un prueba de las carreteras rurales de interés nacional en la provincia de Villa Clara, localizada en la región central de Cuba, con el propósito de analizar la seguridad vial en los tramos que componen dicha red. Para ello, los autores caracterizaron los accidentes vehiculares mediante un procedimiento que contempla cinco aspectos e incluye la determinación de un orden de peligrosidad. Además, determinaron la influencia de cada uno de los elementos componentes del sistema de seguridad vial en la accidentalidad y su interrelación. El área de estudio fue la provincia de Villa Clara, para la cual se particularizó en el aporte a la inseguridad vial, de los elementos relativos a la carretera, a partir del análisis de los modelos de reportes de accidentes de tránsito en tres momentos, contribuyendo de esta forma con medidas, regulaciones o acciones de conservación a favor de la disminución de los accidentes de tránsito, como resultado de un mayor conocimiento de sus causas.

Martínez (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 el prueba, consideró la base de datos de accidentes ocurridos en España durante el año 2012 proporcionada por la Dirección General de Tráfico. Dicha base de datos estaba conformada por 83.115 registros de accidentes con víctimas notificados por los diferentes cuerpos de policía y guardia civil y 202.804 registros referentes a datos específicos de las propias víctimas.

www.bdigital.ula.ve

CAPÍTULO II. MARCO TEÓRICO

1.1. Introducción

La minería de datos ha aumentado su presencia en los últimos años por ofrecer en un sinnúmero 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 pertenece. En la búsqueda de ese objetivo, una organización pública como INPRADEM, con una trayectoria de compromiso con la población, se beneficiará de este proyecto 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. En particular, este proyecto se embarca en el proceso de descripción que ofrece la minería de datos, específicamente la caracterización, que resulta, en este caso en un insumo importante en el proceso de toma de decisiones de la organización con el objeto de mejorar el desempeño del ente.

Desde una perspectiva investigativa la caracterización es una fase descriptiva con fines de identificación, entre otros aspectos, de los componentes, acontecimientos (cronología e hitos), actores, procesos y contexto de una experiencia, un hecho o un proceso (Sánchez, 2010).

La caracterización es un tipo de descripción cualitativa que puede recurrir a datos o a lo cuantitativo con el fin de profundizar el conocimiento sobre algo. Para cualificar ese algo previamente se deben identificar y organizar los datos; y a partir de ellos, describir (caracterizar) de una forma estructurada; y posteriormente, establecer su significado (sistematizar de forma crítica) (Bonilla, Hurtado y Jaramillo, 2009).

En este capítulo se describen algunos conceptos básicos e información necesaria para realizar la caracterización de los registros de emergencias atendidos por INPRADEM.

1.2. Organización

La organización se define como la identificación, clasificación de actividades requeridas, conjunto de actividades necesarias para alcanzar objetivos, asignación de un grupo de actividades a un administrador con poder de autoridad, delegación, coordinación y estructura organizacional, según Koontz y Weihrich (1999).

1.2.1 Organización pública

Las organizaciones públicas son empresas establecidas por el gobierno para proporcionar servicios públicos. Son aquellos entes que incumben al Estado, poseen personalidad jurídica, patrimonio y sistema jurídico propio. Se implantan por razón de un decreto del Ejecutivo, para la actuación en actividades mercantiles, industriales y de servicio o cualquier otra actividad conforme a su denominación y forma jurídica. (Uzcátegui, 2009).

Las organizaciones públicas buscan el bienestar de toda la comunidad mediante la implementación de leyes y/o servicios en pro del bien de la población, asegurando la salud física y mental y garantizando la seguridad de la vida de todos sus habitantes.

1.2.2 INPRADEM

1.2.2.1. Reseña histórica

El 03 de octubre de 1996, la Gobernación del Estado Mérida, pensando en brindarle a Mérida una atención especializada en materia de emergencias, debido a las condiciones geográficas y geomorfológicos que posee el Estado, tomó la iniciativa de crear, por decreto, la Fundación para el Manejo de Emergencias, Desastres Naturales y Defensa Civil del Estado Mérida (FUNDEM).

El 03 de Febrero de 1999 FUNDEM recibe el nombre de Instituto que asume las Funciones de Emergencias, Desastres Naturales y Defensa Civil del Estado Mérida. Es así como se conceptualizó dentro de los ciudadanos en casos de emergencias, diseñando para ello programas y estrategias de atención y prevención. Desde entonces FUNDEM venía trabajando por un solo objetivo, SERVIR y AYUDAR.

Partiendo desde este principio, quienes organizaron las directrices que debían regir a FUNDEM, diseñaron una serie de proyectos y programas a favor del pueblo de Mérida, los

cuales, hoy día pueden ser disfrutados por propios visitantes. Todo esto canalizado a través del Centro de Comunicaciones del Estado Mérida (CECOM) y la Red Teleinformática del Estado Mérida (RETIEM), integrado por todos los organismos responsables de la seguridad a través de su sistema adquirido por la Fundación, para tal fin.

El 01 de Octubre del 2001 se crea el instituto de Protección Civil y Administración de Desastres del Estado Mérida (INPRADEM), ante la eminente situación de riesgo natural y elevada incidencia de accidentes antrópicos que caracteriza el Estado, este organismo es de carácter público, social y humanitario. Tiene su ubicación geográfica en el Estado Mérida, sede principal Av. Los Próceres, Sector Santa Bárbara Oeste.

1.2.2.2. Visión

Ser una organización preeminente en el Estado y de reconocida trayectoria nacional, que garantice altos niveles de seguridad ciudadana en los sectores altamente innovadores en materia de manejo de situaciones de emergencias, desastres naturales y accidentes antrópicos, contribuyendo con la sostenibilidad del desarrollo del Estado Mérida.

1.2.2.3. Misión

Ejecutar las acciones que implican la gestión del riesgo y la minimización de los efectos de los desastres y/o emergencias 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.

2.2.2.4 Objetivo

El Instituto de Protección Civil y Administración de Desastres del Estado Mérida (INPRADEM) tendrá como objeto proteger a la persona y a la sociedad ante la eventualidad de un desastre provocado por agentes naturales o humanos a través de acciones que reduzcan o eliminen la pérdida de vidas, la afectación de la planta productiva, la destrucción de bienes materiales y el daño a la naturaleza, así como asegurar el funcionamiento de los servicios públicos vitales, en coordinación con las dependencias de los tres órdenes de gobierno; de igual modo, administrará los recursos públicos y privados

orientados a minimizar los efectos de los desastres. Trabajar en pro de la población mediante la atención de emergencias que se presente en el Estado Mérida, colaborando con los demás organismo de atención de emergencias para brindar un servicio de calidad a la población.

2.3. Minería de Datos

2.3.1. Definición

La tarea no trivial de extraer información implícita, previamente desconocida y potencialmente útil de bases de datos (Frawley, 1992).

El proceso de descubrir conocimiento interesante de grandes cantidades de datos almacenadas en bases de datos, data warehouses u otro repositorio de información (Han y Kamber, 2001).

2.3.2. Áreas de aplicación

Algunas áreas de aplicación donde hace presencia la minería de datos son:

- Medicina: biomedicina, efectos colaterales de medicamentos, análisis de secuencias genéticas y predicción.
- Finanzas: aprobación de créditos, predicción de quiebras, predicción en el mercado de valores, detección de fraudes, detección de acceso no autorizado a datos de crédito.
- Agricultura: clasificación de enfermedades en tomates y soya.
- Social: datos demográficos, tendencias electorales.
- Mercadotecnia y ventas: identificación de sub-grupos socioeconómicos que muestran comportamiento inusual, patrones de compra, análisis de productos.
- Publicidad: ediciones personalizadas.
- Minería web, minería de blogs, minería de opinión, minería de sentimientos.

2.3.3. Objetivos de la Minería de Datos

La minería de datos persigue dos objetivos de manera general, según (Carrillo, 2007):

- Descripción: el principal producto del proceso de la minería de datos es el descubrimiento de reglas. Estas mostrarán nuevas relaciones entre las variables o excepciones de acuerdo a la empresa en que se utilice este proceso, lo cual enriquecerá el análisis y la descripción de la forma en que trabaja una compañía y ayudará en la planificación y en el diseño de futuros cambios.
- Predicción: una vez descubiertas reglas importantes, estas pueden ser utilizadas para estimar algunas variables de salida. Puede ser en el caso de secuencias en el tiempo, o bien; en la identificación e interrupción a tiempo de una futura mala experiencia de crédito. En esta tarea, se complementan las técnicas estadísticas tradicionales con aquellas provenientes de la inteligencia artificial.

2.3.4. Metodologías

Algunas de las metodologías más conocidas en esta área son (León, 2017):

- **SEMMA (Sample, Explore, Modify, Model, Assess):** fue propuesta por SAS Institute Inc y definida como proceso de selección, exploración y modelamiento de grandes cantidades de datos para descubrir patrones de negocios desconocidos. Las fases que contempla se muestran en la figura 1:

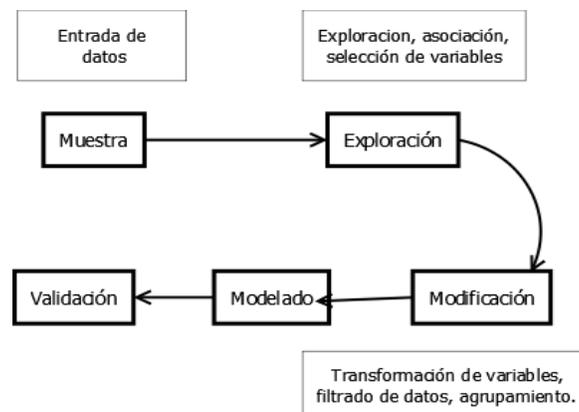


Figura 1. Fases de la metodología SEMMA

- **KDD (Knowledge Discovery in Databases):** es una metodología propuesta por Fayyad en 1996, constituida por cinco fases: selección, pre procesamiento, transformación, minería de datos y evaluación e implantación. Es un proceso iterativo e interactivo como se observa en la figura 2.

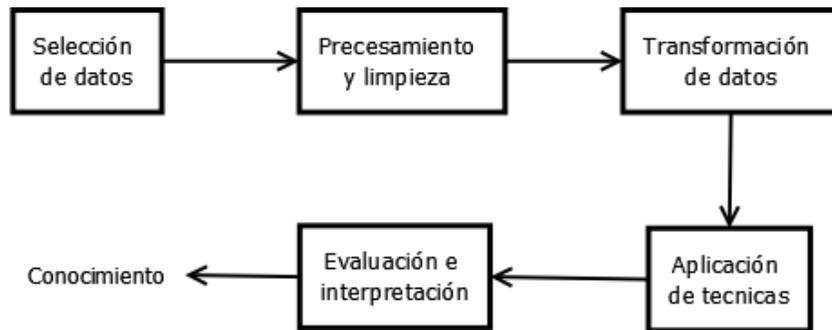


Figura 2. Fases de la metodología KDD

- **CRISP-DM (Cross-Industry Standard Process for Data Mining):** es una de las metodologías más usadas en minería de datos. En la figura 3 se muestra una visión general de las fases.

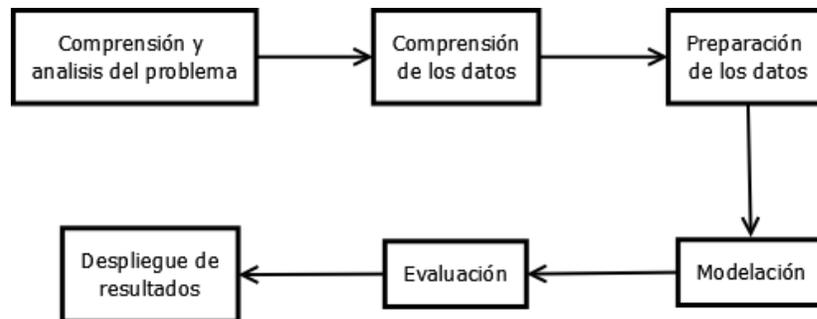


Figura 3. Visión general CRISP-DM

2.4. Metodología CRISP-DM

CRISP-DM, creada por el grupo de empresas SPSS, NCR y Daimler Chrysler en el año 2000, es actualmente la guía de referencia más utilizada en el desarrollo de proyectos de minería de datos. Estructura el proceso en seis fases. La sucesión de fases, no es necesariamente rígida. Cada fase es descompuesta en varias tareas generales de segundo nivel. Las tareas generales se proyectan a tareas específicas, pero en ningún momento se propone como realizarlas, es decir, CRISP-DM establece un conjunto de tareas y actividades para cada fase del proyecto pero no especifica cómo llevarlas a cabo. Estas actividades dependerán del propósito del proyecto que se desarrolle (Gordillo y Haedo, 2011).

CRISP-DM, es el método que se adapta mejor para el desarrollo de este proyecto, y por ende, se tomará como referencia en este proyecto. En la figura 4 se detallan las fases que comprenden dicha metodología:

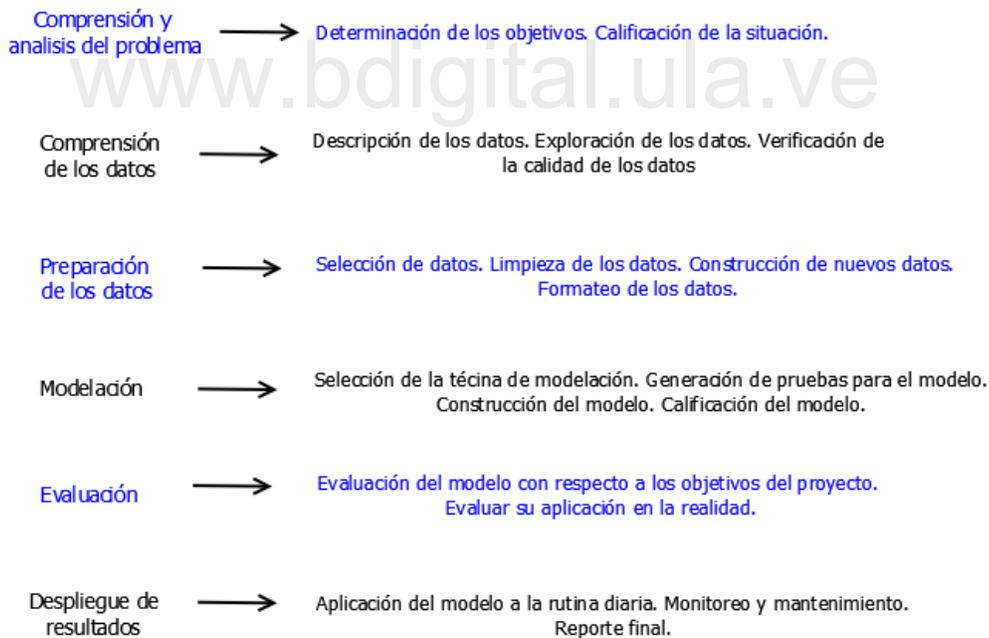


Figura 4. Fases de la Metodología CRISP-DM

Tomando como referencia a Chapman (2000) las fases de la metodología CRISP-DM se explican a continuación:

1. **Análisis del problema:** incluye la comprensión de los objetivos y requerimientos del proyecto, con el fin de convertirlos en objetivos técnicos y en una planificación, pesando en el cómo abordar el problema y las diferentes opciones con las que se podrían trabajar.
2. **Comprensión de datos:** comprende la recolección inicial de datos, de modo que sea posible establecer un primer contacto con el problema, identificando la calidad de los datos y estableciendo las relaciones más evidentes que permitan establecer las primeras hipótesis. En esta fase se debe analizar y predefinir la manera en se abordará el problema en materia de software e incluso técnicas, ya que la siguiente fases dependerá de estas decisiones.
3. **Preparación de datos:** incluye las tareas generales de selección de datos a los que se va a aplicar la técnica de modelado (variables y muestras), limpieza de los datos, generación de variables adicionales, integración de diferentes orígenes de datos y cambios de formato. Esta fase se encuentra relacionada con la fase de modelado, puesto que en función de la técnica de modelado que vaya a ser utilizada, los datos necesitan ser procesados para que puedan ser aplicados a ese software y técnicas en particular. Por lo tanto, las fases de preparación y modelado interactúan de forma sistemática.
4. **Modelado:** se seleccionan de manera definitiva las técnicas de modelado más apropiadas para el proyecto en función de los siguientes criterios: ser apropiada al problema, disponer de datos adecuados y conocimiento de la técnica. Los parámetros utilizados en la generación del modelo dependen de las características de los datos.
5. **Evaluación:** se evalúa el modelo, no desde el punto de vista de los datos, sino desde el cumplimiento de los criterios de éxito del problema. Se debe revisar el proceso seguido, teniendo en cuenta los resultados obtenidos para poder repetir algún paso en el que, a la vista del desarrollo posterior del proceso, se hayan podido cometer errores. Si el modelo generado es válido en función de los criterios de éxito establecidos en la primera fase, se procede a la aplicación del modelo.

Normalmente los proyectos de explotación de información no terminan en la implantación del modelo, sino que se deben documentar y presentar los resultados de manera comprensible en orden a lograr un incremento del conocimiento. Además en la fase de explotación se debe de asegurar el mantenimiento de la aplicación y la posible difusión de los resultados.

2.5. Minería de datos en las organizaciones

La minería de datos es conocida por ser usada en diferentes organizaciones con diversos fines, para la extracción de datos y su análisis mediante técnicas estadísticas de grandes bases de datos brindando conocimiento que le sea de utilidad que lleve a mejorar algún aspecto de importancia para la organización. Por ejemplo, en las empresas es ampliamente usada para prospectar clientes, para saber quiénes son los mejores candidatos para un programa o un producto en especial, para analizar el producto o parte de un proceso de producción.

El tema de minería de datos está ligado al servicio, aporta respuestas a lo que está pasando en mercadeo y servicio, es una puerta para las peticiones, quejas y reclamos, si se trabaja adecuadamente esta información, se pueden establecer cambios en servicios y mejorar las operaciones de la empresa.

2.5.1. Minería de datos en los procesos de toma de decisiones

En la actualidad las organizaciones son impulsadas por el deseo de un cambio continuo; específicamente una mejora continua, por ello, las empresas privadas tanto como las públicas deben tener la capacidad de ser adaptativas, flexibles y abiertas a cambios constantes, contando con la facilidad para aprender cómo resolver problemas y generar conocimiento, y así poder establecer nuevos métodos en pro de la resolución de los mismos.

En las organizaciones la necesidad de información surge como consecuencia de dos factores: la incertidumbre y el coste a asumir en caso de errar en la toma de decisiones (Wright y Ashill, 1998). Actualmente a raíz de estos factores, la información es de gran

importancia haciendo presencia de manera permanente en la búsqueda para alcanzar el objetivo previsto, hasta el punto de que algunos autores hablan de una economía basada en la información y el conocimiento. En este sentido, algunas empresas han adoptado la política de recopilación de datos como parte imprescindible en su desarrollo.

El interés del mundo profesional por la minería de datos se ve reflejado en el variado y amplio conjunto de empresas que han lanzado al mercado productos de *data mining*, entre las cuales se encuentran las tres empresas más importantes en herramientas de análisis: SPSS (Clementine), SAS (Enterprise Miner) e IBM (Intelligent Miner). Desde el punto de vista de los expertos se habla de la minería de datos como uno de los campos emergentes y con gran futuro en la investigación de *marketing*.

Todas estas pruebas, implementación y la manera en que se ha expuesto la minería de datos, han incrementado el deseo desenfrenado por demandar un mayor control de los procesos u operaciones y servicios, que la minería de datos promete poner a disposición y dando como resultado inminente un núcleo para una gestión global y fundamental, proporcionando servicios de calidad y logrando un rendimiento óptimo de las inversiones, en infraestructuras, en un entorno dirigido hacia una gestión de demandas.

Es importante resaltar que poseer datos no es el elemento esencial en una toma de decisión acertada; es extraer de los datos información que al ser analizada resulte en conocimiento que pueda ser usado para la implementación, esto es lo que proporciona el apoyo necesario para la toma de una decisión argumentada, que oriente a la empresa hacia el cumplimiento de sus metas y objetivos. El nivel de satisfacción del cumplimiento del objetivo de este proyecto se mide a través de qué tan valioso es el conocimiento que se va a obtener a partir de la extracción y análisis de información, es decir, en qué medida va a ayudar este conocimiento a alcanzar las metas de INPRADEM, las cuales están centradas a la calidad del servicio que brinda a la población del Estado Mérida.

CAPÍTULO III. MARCO METODOLÓGICO

3.1. Introducción

En este capítulo se describe el proceso de caracterización de los datos registrados de las emergencias atendidas por INPRADEM, tomando como guía la metodología CRISP-DM explicada anteriormente. Para ello, se establecen esquemas de representación de los datos para que sean de provecho en las pruebas que se realizan, implementando estudios que permitan llevar a una caracterización apropiada. El tratamiento de los datos se realizó en Excel de Microsoft Office 2013 y RStudio.

3.2. Fases 1 y 2 de la Metodología CRISP-DM

3.2.1. Compresión 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.

3.2.2. Entendimiento de 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. Los datos corresponden a 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

la tabla 1 se presenta una breve descripción de los datos. En los anexos en digital puede verse en su totalidad el conjunto de datos considerados en el prueba.

Tabla 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. Ej.: 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 hora 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. Ej. 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. Ej. Eliminación de riesgo. Traslado de emergencia

3.3. Fase 3: Preparación de los datos

En esta fase se procedió a la selección y limpieza de los datos, partiendo de la base de datos que se muestra en la tabla 2.

Tabla 2. Bases de datos inicial

SECTOR	META PROGRAMADA	META ALCANZADA	% DE CUMPLIMIENTO	DEPARTAMENTO , FUNCIONARIO RESPONSABLE DE LA ACTIVIDAD	NOTAS EXPLICATIVAS:	CONSEJOS COMUNALES O CUALQUIER OTRA ORGANIZACIÓN ABORDADA
Mucuchies	1	1	100%	Oficial Sar: Mariano Villarreal		
El Ro,al	1	1	100%	Oficial Sar: Albert Quintero		
Mucuchies	1	1	100%	Oficial Sar: Luis Rojas		
Moconoque	1	1	100%	Oficial Sar: Mariano Villarreal		

Al analizar la tabla 2 se observa que las siguientes variables no aportan información útil para el prueba, por tanto, se eliminan:

- Horas hombre: esta variable indica el número de horas que acumulan entre todos los funcionarios que participan en el evento, es un dato que se puede obtener de las variables: número de funcionario y total de horas-
- Denominación producto meta: definida dentro de número de producto meta y actividades, por lo que se considera que tomando en cuenta las otras dos variables es suficiente para determinar y clasificar el tipo de actividad que se realiza.
- Cantidad de actividades ejecutadas: en su mayoría la tabla presenta siempre un número entero, uno o simplemente se encuentra en blanco por esta razón se excluye.
- Descripción de la actividad: texto libre que entra en detalle sobre la actividad realizada es difícil extraer información útil de este campo.
- Coordenadas UTM, sector: brindan el punto exacto en donde se realiza la actividad. Existen otras variables que engloban de mejor manera el lugar. Además, este campo se encuentra en blanco para muchas actividades.
- Población beneficiada: este campo solo posee como valor único el número uno por lo tanto también se descarta.
- Meta programada, meta alcanzada y porcentaje de rendimientos: el primero solo indica si la actividad es planificada o no, por ejemplo: una charla. La segunda, indica que la meta ha sido alcanzada al momento de cubrir la actividad y el tercero siempre tiene un porcentaje 100% de rendimiento.

- Notas explicativas: en su mayoría se encuentra en blanco al menos que al momento de explicar en detalle la actividad haya quedado algo por fuera digno de mención, y es también un campo de texto libre.
- Consejos comunales: este campo en su mayoría se encuentra en blanco al menos que sea alguna actividad en específico que se esté desarrollando como guardia de prevención, a pesar de eso, el consejo comunal que indicaría el sector o parroquia se puede apreciar en otra variable que aporta la tabla.
- Cantidad de material entregado y medio de verificación: la primera solo se utiliza cuando se realizan avalúos de terreno o edificación, es decir, unas actividades específicas, el resto de las actividades poseen este campo en blanco. El medio de verificación también se llena solo para ciertas actividades, por ejemplo: algún lesionado que fue trasladado a un centro médico, en este caso, el medio de verificación sería una planilla de atención integral. Sin embargo, son muy puntuales esas actividades por eso se excluye.
- Número de libro y número de página: ambos campos son referencias al lugar en donde se encuentra la planilla en físico de esa actividad, ya que en realidad para el prueba de los datos no interesa donde este guardada la referencia a esa actividad se excluyen. Estos campos solo fueron usados para contrastar las planillas que fueron otorgadas en digital con las físicas y corroborar los datos. Para este proceso solo se seleccionaron algunas planillas de cada centro y se compararon con los datos en su formato en digital.

En la tabla 3 se presentan las variables que fueron seleccionadas para la prueba:

Semana, Día, Hora de Inicio, Hora final, Total de horas efectivas, Cantidad de funcionarios, Numero de producto meta, Actividad, Municipio, Departamento, Equipos y Materiales utilizados.

Tabla 3. Datos después del primer filtrado

SEMANA	FECHA	HORA INICIO	HORA FINAL	TOTAL HORAS EFECTIVAS	CANTIDAD DE FUNCIONARIOS	HORAS HOMBRE
Semana_01	miércoles, 01 Enero 2014	20:00	22:00	02:00	3	6:00:00
Semana_01	miércoles, 01 Enero 2014	21:55	22:25	00:30	4	2:00:00
Semana_01	jueves, 02 Enero 2014	13:20	16:30	03:10	3	9:30:00
Semana_01	jueves, 02 Enero 2014	19:10	20:10	01:00	4	4:00:00

Con estas once (11) variables de entrada se realiza un cambio a las variables tipo texto para llevar todos los datos a carácter numeral, asignándoles un código (número) a cada uno. En las tablas 4 y 5 se presenta un ejemplo con los cambios realizados para las variables Día y Municipio, respectivamente.

Tabla 4. Ejemplos de variables llevadas a su código numeral

Lunes	1
Martes	2
Miércoles	3
Jueves	4
Viernes	5
Sábado	6
Domingo	7

Alberto Adriani	1	Miranda	13
Antonio Pinto Salinas	2	Obispo Ramos de Lora	14
Andrés Bello	3	Padre Noguera	15
Aricagua	4	Pueblo Llano	16
Arzobispo Chacón	5	Rangel	17
Campo Elías	6	Rivas Dávila	18
Caracciolo Parra Olmedo	7	Santos Marquina	19
Cardenal Quintero	8	Sucre	20
Guaraque	9	Tovar	21
Julio cesar Salas	10	Tulio Febres Cordero	22
Justo Briceño	11	Zea	23
Libertador	12		

Luego de este cambio la base de datos con la que se va a trabajar es toda numérica y corresponden a un código asignado a cada campo (ver tabla 5).

Tabla 5. Datos después de asignarle su código numeral

SEMANA	DIA	HORA INICIO	HORA FINAL	TOTAL HORAS	N FUNCIONARIOS	N PM	ACTIVIDAD	DEPARTAMENTO	EQUIPO	MUNICIPIO
1	3	1	3	2	3	10	8	4	1	17
1	3	3	3	1	4	10	13	4	3	17
1	3	4	4	1	3	10	22	10	1	17
1	7	4	4	1,5	4	10	12	11	3	17
1	6	4	4	2	3	10	8	3	3	17
1	6	4	4	2	3	10	8	3	3	19
1	4	4	4	1	3	10	11	10	11	17
1	4	4	1	4	3	10	23	3	9	17
1	4	4	1	4	3	10	23	3	9	17
1	3	4	4	1	3	10	8	4	11	17
1	6	4	4	1	3	7	8	4	3	13
1	6	4	4	1	2	10	22	11	1	17
1	1	4	4	1	2	10	8	9	2	12
1	5	4	4	1	2	10	8	4	1	12
1	4	4	4	1,5	4	10	12	9	3	12
1	4	4	4	1,5	4	10	12	9	3	12
1	7	4	4	2	10	11	16	9	3	12
1	5	4	2	4	3	10	13	4	1	19

Finalmente, se procede a realizar el estudio descriptivo mediante gráficos de histogramas, análisis de frecuencia y de correlación utilizando RStudio.

El análisis de correlación no arrojó relación entre las variables. Por tanto, se decide realizar un nuevo filtrado en el cual se eliminará la variable “Producto Meta” por ser una que engloba una serie de actividades que valen la pena estudiar individualmente. Al realizar nuevamente el análisis de correlación se observa que todos los valores son muy cercanos al cero, dando a entender que las variables son independientes entre sí. En tal sentido, no es posible hacer una reducción de variables tomando en cuenta su correlación lineal.

Antes de proceder a realizar la primera prueba, basada en gráficos de frecuencia, es importante ilustrar en un esquema cómo se pueden clasificar las variables con las que se trabajará de ahora en adelante (ver Figura 5).

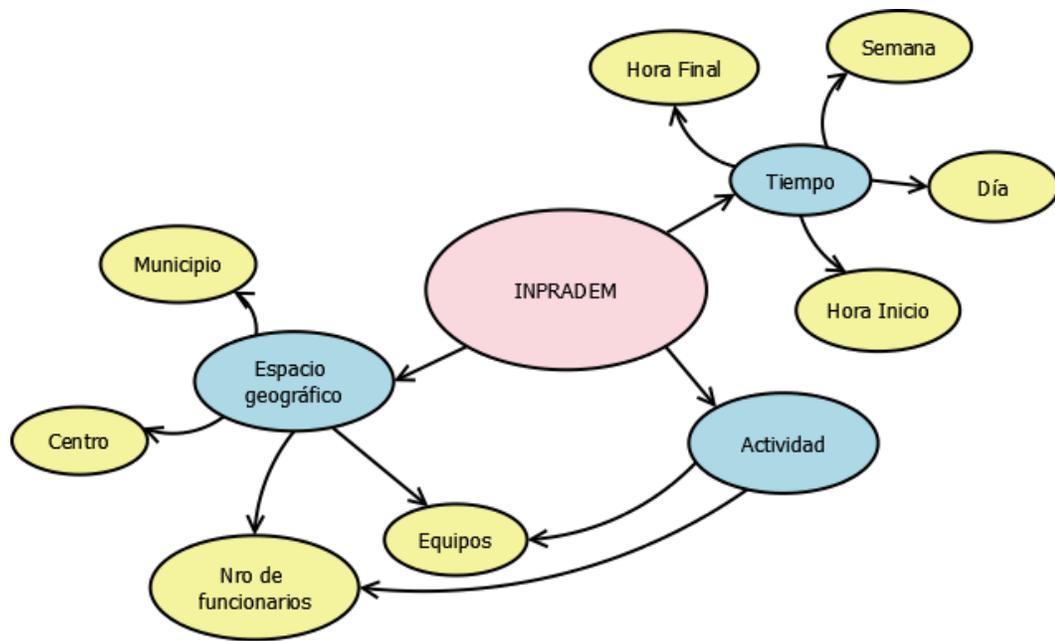


Figura 5. Esquema de clasificación de las variables

Este esquema permite observar que las variables se pueden clasificar en tres grandes grupos:

- Aquellas que están relacionadas con el tiempo, ya sean intervalos de tiempo como Semana y día, así como horas específicas: hora de inicio y final.
- La variable Actividad se encuentra sola ya que no tiene relación alguna con el tiempo ni el espacio geográfico en que se desarrolle, pero ella sí influye en otras variables.
- Por último, se encuentra 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 depende 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.

3.4. Fase 4: Pruebas iniciales

3.4.1. Gráfico de frecuencia para cada variable

Debido al carácter independiente de las variables se procedió a realizar una primera prueba por variable basado en las frecuencias máximas. Para cada variable se mostrará una lista en orden ascendente de las frecuencias máximas:

- **Semana:** varían a lo largo del año pero los meses de noviembre y septiembre son aquellos en los que más se prestan servicios durante el año, ya que varias semanas pertenecientes a estos meses hacen presencia en la tabla 6 donde se muestran las semanas con mayor demanda.

Tabla 6. Variable semana

Semana	Mes del año	Nro. de eventos
45	Primeros días de noviembre	703
9	Últimos días de febrero	684
37	Mediados de septiembre	682
46	Mediados de noviembre	680
2	Segunda semana de enero	672
38	Finales de septiembre	671
39	Última semana de septiembre	666

- **Día:** Jueves es el día que presenta más actividades, seguido por miércoles y martes. (ver tabla 7).

Tabla 7. Variable día

Código	Día	Nro. de eventos
4	Jueves	4446
3	Miércoles	4419
2	Martes	4241
1	Lunes	4170
5	Viernes	4011
6	Sábado	3619
7	Domingo	3026

- **Hora Inicio:** se observa que la mayoría de eventos se desarrollan en la mañana (ocho, nueve y diez de la mañana son las horas con mayor frecuencia) (ver Tabla 8).

Tabla 8. Variable Hora de Inicio

Código	Horas	Nro. de eventos
8	8:00 am	3740
9	9:00 am	3012
10	10:00 am	3010
14	2:00 pm	2445
11	11:00 am	2416
15	3:00 pm	2121

- **Hora Final:** en la tarde finalizan la mayoría de las actividades (ver Tabla 9).

Tabla 9. Variable Hora Final

Código	Horas	Nro. de eventos
18	6:00 pm	3053
11	11:00 am	2789
10	10:00 am	2515
16	4:00 pm	2312
12	12:00 pm	2302
15	3:00 pm	2217

- **Número de funcionarios:** uno, dos y tres es el número de funcionarios que se despachan para cubrir las actividades en su mayoría (ver Tabla 10).

Tabla 10. Variable Nro. de Funcionarios

Nro. de funcionarios	Nro. de eventos
1	11229
2	8344
3	4916
4	2053
5	843

- **Actividad:** Las actividades con mayor frecuencia se presentan en la tabla 11, en orden descendente según su número de eventos.

Tabla 11. Variable Actividad

Código de la actividades	Descripción	Nro. De eventos
15	Guardia de prevención	4292
19	Evaluaciones de riesgo de terreno	3774
9	Evaluaciones de riesgo de edificación	3006
13	Traslado regular	2796
25	Evaluaciones de análisis de riesgo	1924
11	Atención en el sitio	1603
8	Traslado de emergencia	1514
7	Eliminación de riesgo	1398
12	Atención hechos viales	1194
10	Recorrido vial	1143

Las cinco (05) actividades de mayor frecuencia que se desarrollan en la organización son eventos que requieren planeación previa, de sexto se encuentra atención en el sitio que es la primera actividad no planeada que se ve en la lista.

- **Departamento:** en la tabla 12 se presentan los departamentos que prestan un mayor servicio dentro de la organización.

Tabla 12. Variable Departamento

Código	Departamento	Nro. de eventos
3	SAR I	13841
9	Sin departamento	4146
4	SAR	3719
12	Respuesta inmediata	1202

- **Equipo:** los equipos más usados para prestar servicios dentro de la organización se presentan en la tabla 13.

Tabla 13. Variable Equipo

Código	Equipo	Descripción	Nro. de eventos
3	SAR	Se compone de camionetas pick-up o rústicos	7022
10	Sin Unidad	Solo se despachan oficiales	4301
1	ROJO	Ambulancias	3955
8	Unidad particular	Uso de vehículo particular	3307
7	Material de Oficina	Planilla de evaluaciones de riesgo	3088
14	Primeros auxilios	Kit de primeros auxilios	1885
5	HALCON	Flota de motos	1368

- **Municipio:** en la tabla 14 se listan los municipios que más eventualidades presentaron durante los años bajo estudio.

Tabla 14. Variable Municipio

Código	Municipio	Nro. de eventos
21	Tovar	4701
12	Libertador	3806
1	Alberto Adriani	3308
17	Rangel	3011

En la figura 6 se presentan las frecuencias para las variables mencionadas previamente.

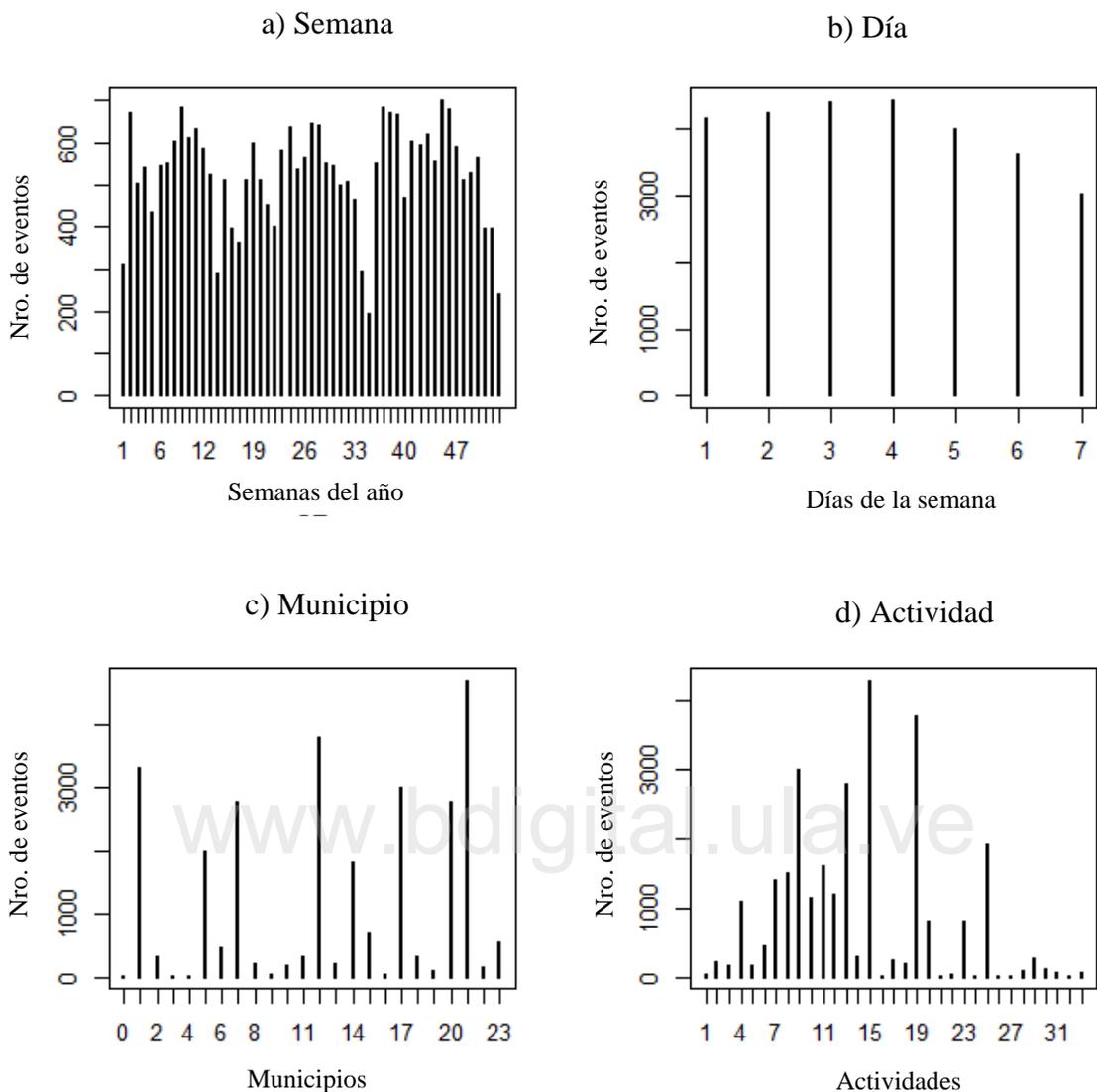


Figura 6. Frecuencias para las variables Semana, Día, Municipio y Actividad

Los resultados obtenidos de este primer estudio arrojan como información relevante que la organización desarrolla en su mayoría eventos que requieren planeación previa, en la tabla 6 se aprecian las cinco primeras actividades seguidas por atención en el sitio de sexto lugar, luego traslado de emergencia de octavo y atención a hechos viales de doceavo. Esto permite detectar que las actividades que se desarrollan se pueden dividir entre actividades planeadas y no planeadas.

En tal sentido, es importante filtrar más los datos antes de proceder con los otros análisis, ya que son necesarios estudios adicionales como las horas en las que se dan las actividades en la organización, para tener una mejor idea del camino a tomar en las pruebas siguientes.

3.4.2. Por día las horas en que se presentan los eventos

En su mayoría los eventos se presentan en la mañana, específicamente, entre las ocho y nueve de la mañana, según los resultados de la prueba del primer estudio. Ahora, se decide analizar para cada día de la semana la hora en que se despachan las unidades para atender las diferentes actividades. En la figura 7 se presentan los gráficos obtenidos para los primeros cuatro días (lunes hasta el día jueves). Los demás gráficos se encuentran en el anexo digital.

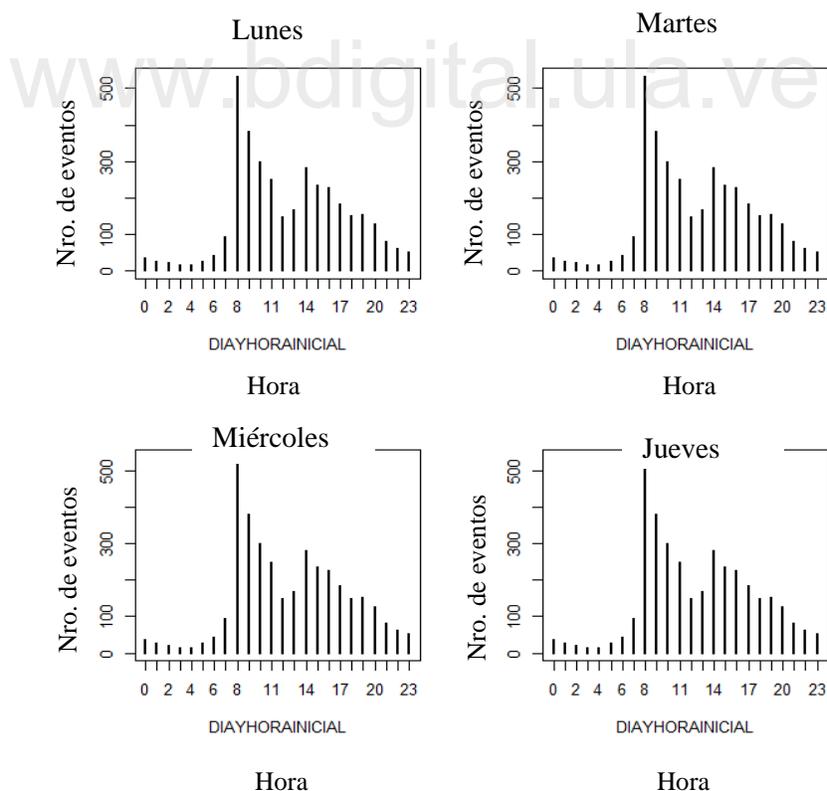


Figura 7. Hora en que presentan los eventos para los días lunes, martes, miércoles y jueves

Esta prueba confirma que las actividades comienzan aproximadamente a las ocho de la mañana, lo cual no tiene sentido tomando en cuenta que INPRADEM atiende a la población en casos de emergencia; este análisis junto con el resultado obtenido en la prueba relaciona a la variable Actividad en el primer estudio realizado (sección 3.4.1) en donde se observa que la mayoría de las actividades que se realizan son las que llevan una planeación previa, lleva a suponer que la segunda prueba está siendo afectada por esas actividades. En tal sentido, se decidió realizar una nueva prueba tomando del análisis anterior las actividades que poseen mayor frecuencia dentro de la organización y estudiar cada una de manera independiente, observando el comportamiento de las diferentes variables con respecto a cada actividad, para encontrar alguna información relevante que permita realizar un filtrado que encauce el proyecto en dirección provechosa.

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

Esta prueba permite apreciar claramente la diferencia que existe entre las actividades que se desarrollan en la organización. En la figura 8 se observa que las actividades como: a) análisis o b) eliminación de riesgo poseen picos altos en horas muy puntuales como ocho, nueve y diez de la mañana, luego disminuyen durante el mediodía y aumenta otra vez durante las primeras horas de la tarde. De igual forma, se observa que estos servicios son casi inexistentes en horas de la madrugada y horas cercanas a la medianoche, contrario al comportamiento de las actividades como atención en el sitio o traslado de emergencia que se comportan muy diferentes durante el transcurso del día, y se presentan durante horas más variadas, aumentando considerablemente al comienzo del llamado horario de oficina (alrededor de las ocho de la mañana) pero se mantiene fluctuando un poco a lo largo del día hasta la noche donde comienza a disminuir ya en un horario cercano a la media noche.

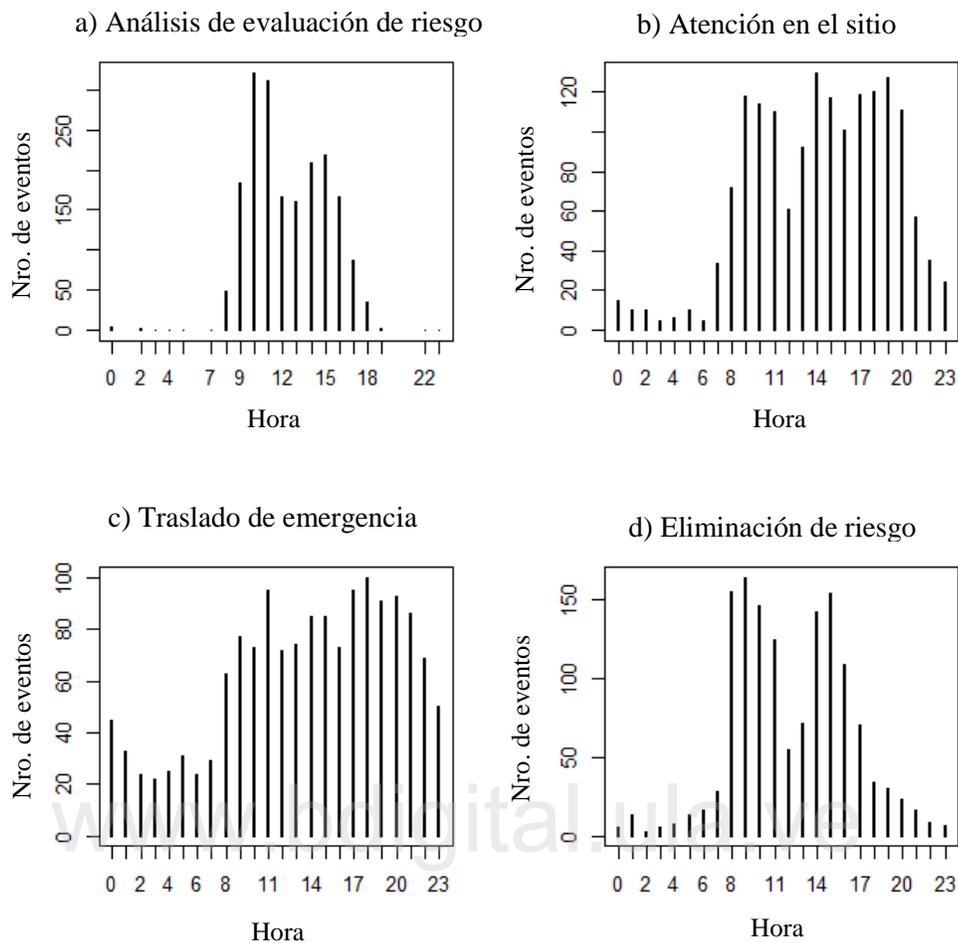


Figura 8. Horas en que se presta el servicio para diferentes actividades

A partir de esta prueba se observó la diferencia existente entre las actividades programadas y no programadas de la organización, apreciándose claramente la diferencia entre las horas en que se desarrollan las actividades programadas (horas de la mañana) y las no programadas que se presentan en horas más variadas.

Tomado de las pruebas anteriores, que las actividades realizadas por INPRADEM pueden ser programadas o no, y después de un análisis se considera que aporta más a la organización y a este proyecto, estudiar las actividades no programadas, es decir, las emergencias que se presentan por su importancia en lograr un mejor manejo y estructuración de la organización para hacerle frente a las emergencias consiguiendo un mejor manejo de recursos, tiempos de respuestas, etc. Para esto, se tomarán las tres actividades no programadas que están incluidas en las diez primeras que se presentan con

más frecuencia en INPRADEM (11 atención en el sitio, 8 traslado de emergencia y 12 hechos viales). Estas tres actividades se estudian con respecto a las demás variables, como prueba inicial.

3.5. Pruebas entorno a las actividades no programadas

El Estado Mérida cuenta con varias sedes de INPRADEM ubicadas en puntos estratégicos para ofrecer una respuesta inmediata y cercana al lugar de la emergencia. En total son ocho sedes que desempeñan diferentes actividades cubriendo los servicios que necesita la población. Cada centro se encarga de lugares geográficos distintos del Estado, así como algunos comparten responsabilidades para algunos sectores y trabajan en conjunto. Es por ello que para las pruebas relacionadas a las actividades no programadas se decide agregar la variable centro.

3.5.1. Comportamiento general de actividades con respecto a cada variable

En la tabla 15 se describe el comportamiento general de cada variable con sus valores más altos respectivamente

Tabla 15. Comportamiento de las actividades no programadas con respecto a cada variable

VARIABLES	Resultado (cantidad de eventos)
Semana	Segunda semana del año (117), Tercera semana de marzo (117), 2da 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 24horas: 19 (323), 18 (309), 20 (301), 17 (292)
Departamento	SAR (1644), Respuesta Inmediata (580)
Centro	Sur del Lago (856), Panamericana (576) y Libertador (541)

Esta prueba destaca dos variables en particular Centro y Hora de Inicio, estos tres centros en particular atienden una cantidad mayor de emergencia comparado con el resto, por ello, las pruebas siguientes se harán tomando en cuenta estos tres en particular. Así mismo, la prueba resaltó que las horas en las que se presentan las emergencias son en un horario tarde-noche contrario a las pruebas anteriores cuando se tomaban todas las actividades que indicaban que se desarrollaban en un horario matutino.

3.5.2. Comportamiento por año de las actividades con respecto a cada variable

Para esta prueba se agregó la variable año para separar las emergencias suscitadas por año y poder descubrir si existe cambios en los comportamientos de dichas variables en el transcurso de los tres años contenidos en la data bajo estudio (ver Figura 9).

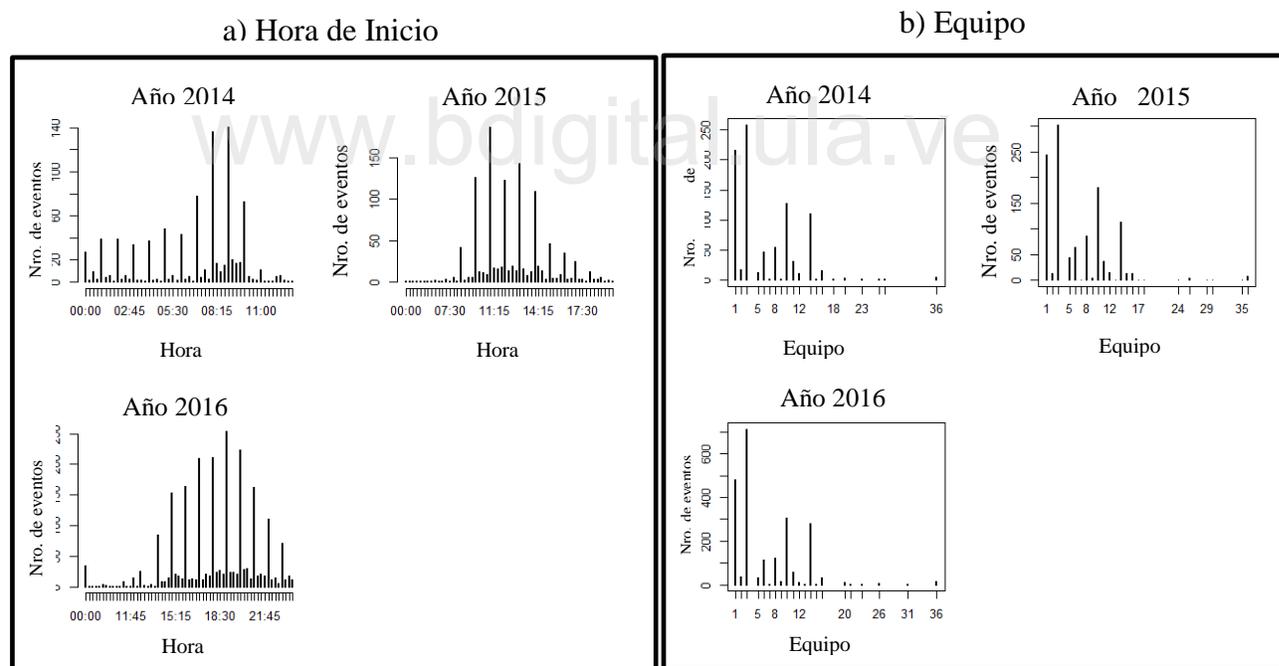


Figura 9. Comportamiento por año de las variables Hora de Inicio y Equipo de las actividades no programadas

En la figura 9 se observa que el comportamiento de las variables varía durante los tres años. Por otra parte, la variable a) hora de inicio cambia significativamente, mientras que, la variable b) equipo presenta pocos cambios. Adicionalmente, se observan horas con

picos bastante altos, así como también equipo y centros con alta demanda, surgiendo una prueba orientada a colisiones.

Una colisión para los pruebas de este proyecto 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, las colisiones son tomadas con una diferencia de quince minutos. Cabe destacar, que 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 10), 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 top que se obtiene antes de dispararse el comportamiento de la curva, por ende este intervalo es el que será usado para las pruebas correspondientes a colisiones.

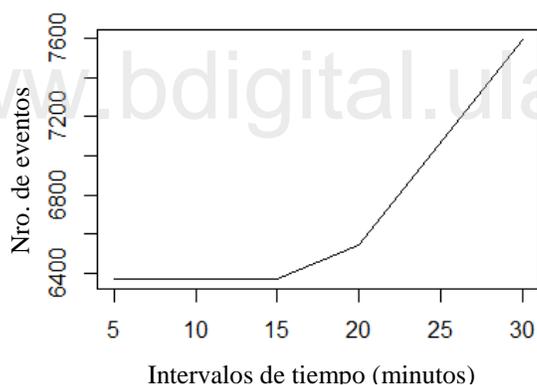


Figura 10. Curva de las colisiones en cada intervalo de tiempo

3.5.3. Prueba comparativa entre el comportamiento de las variables en presencia de colisiones durante los años bajo estudio

Se compara el comportamiento de la prueba anterior (sección 5.3.2) con esta nueva prueba, en donde se obtiene a 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 11 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

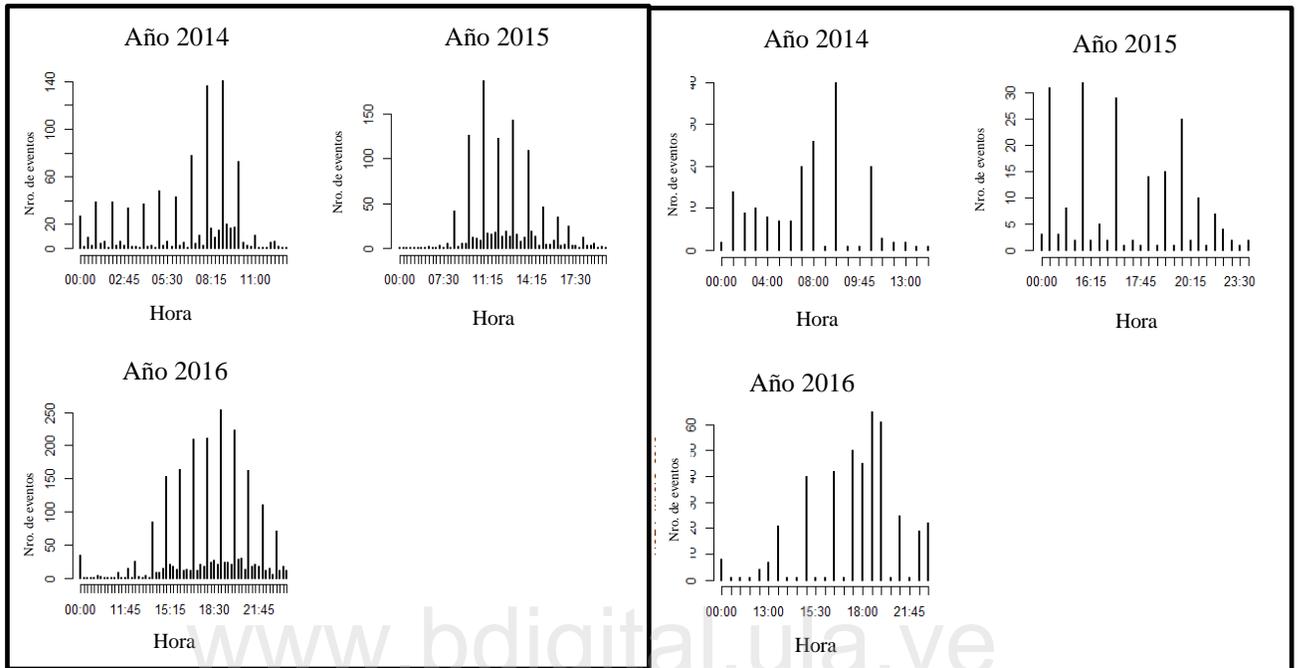


Figura 11. Comparación de la variable hora de inicio para los tres años bajo estudio con o sin prueba de colisión

Se podría suponer que las horas en que se presentan más emergencias son aquellas en las que se espera encontrar mayor número de colisiones. Sin embargo, los resultados indican que el comportamiento de las colisiones presenta algunas diferencias, por ejemplo, si se observa la Figura 11 y se compara cualquier año de a) Hora de inicio sin colisión con el mismo año en b) Hora de inicio con colisión, a pesar de presentar coincidencias para algunas horas, el comportamiento general es diferente. A continuación 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 (ver Tabla 16).

Tabla 16. Tabla comparativa de las variables con o sin prueba de 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	Sabado	Viernes	Jueves	Sabado
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 Mocoties	Panamericana Mocoties Sur del Lago	Sur del Lago Panamericana Libertador	Panamericana Libertador Paramo	Sur del Lago Panamericana Mocoties	Sur del Lago Libertador Panamericana

En la tabla 16, Nro. de semana varia su comportamiento, se observa la presencia de semanas distintas para el mismo año en la atención de emergencias sin o con colisiones. La variable día tiene un comportamiento similar, con la excepción del año 2014 donde el día en el cual se producen más colisiones pasa de ser jueves a viernes. Por su parte, Hora de inicio mantiene su comportamiento para los años 2014 y 2016 mientras que el año 2015 presenta variaciones, sin colisiones las emergencias se presentan en su mayoría en la mañana, mientras que en presencia de colisiones las horas ahora se presentan en la tarde. Para el caso de Equipo, SAR está siempre presente, pero Rojo desaparece cuando hay colisiones. Por último, la variable centro solo varia un poco la cantidad en el número de demandas a los diferentes centros; pero Libertador, Panamericana y Sur del Lago son los centros en los cuales se presentan más demandas y colisiones.

Cabe destacar que la primera prueba realizada tomando las emergencias (sección 3.5.1) permitió descubrir el comportamiento general de las emergencias que se presentan en el Estado Mérida y compararlas a través de los tres años, las diferencias observadas en este prueba entre el comportamiento de las variables en presencia o no de colisiones, guía este proyecto a tomar las colisiones como eje central para una nueva prueba.

3.6. Pruebas entorno a las colisiones

3.6.1. 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. Para ello, se consideró los centros como variable principal (ver Figura 12), tomando los centros en donde se suscitan mayor cantidad de emergencias obtenidos en el prueba de la sección 3.5.1 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.

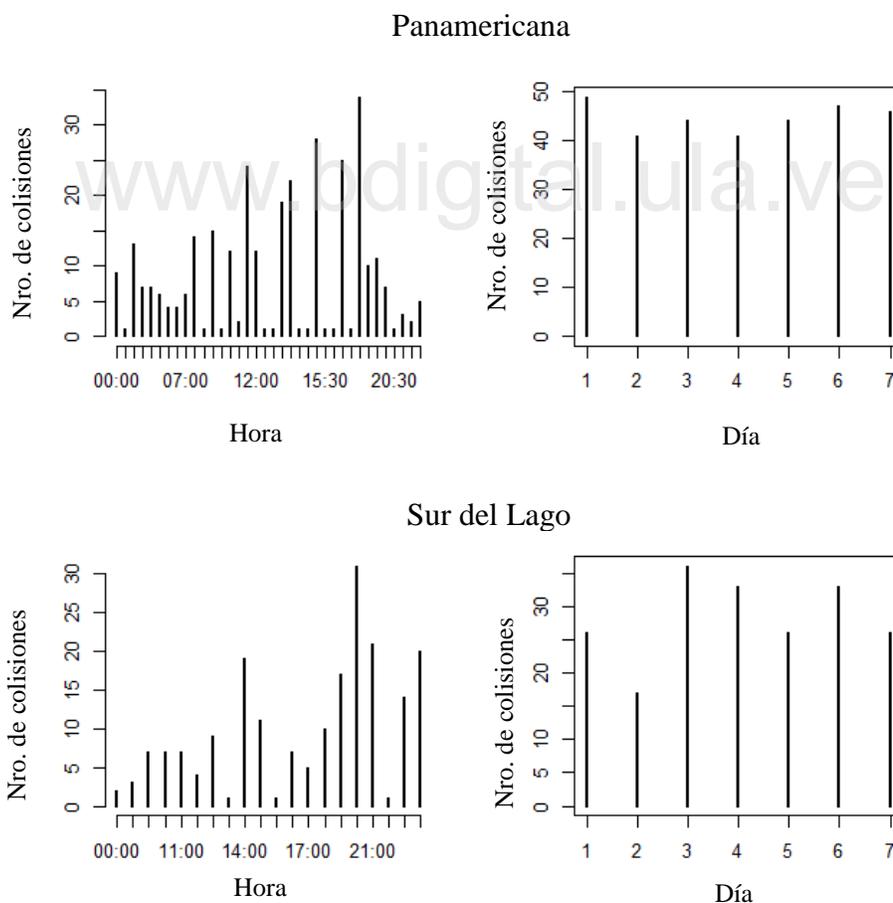


Figura 12. Para los centros Panamericana y Sur del Lago, comportamiento de las colisiones para las variables Hora y Día

Esta prueba permitió descubrir que cada centro se comporta de manera distinta tanto en los días como en las horas en que se presentan las emergencias. En este caso, las colisiones muestran que en efecto cada centro posee una demanda a determinadas horas y días diferentes, analizando el motivo de esta diferencia se podría suponer que se debe al distinto estilo de vida que llevan las personas en cada municipio del Estado.

Las colisiones en horas y días permitieron determinar las fallas que posee esos centros al momento de atender el llamado de emergencia. En tal sentido, es importante crear una nueva prueba para observar en detalle el comportamiento para cada centro.

3.6.2. Horas en que se presentan las colisiones en cada centro durante el transcurso de la semana

En la figura 13 se presentan los resultados, observándose de manera puntual las horas en que se producen las colisiones en el transcurso de la semana para cada centro, separándolo por día, así como retardo en la atención de la emergencia. Dichos resultados pueden reflejarse como puntos de mejoras. Se precisa en esta prueba los puntos, en este caso, 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.

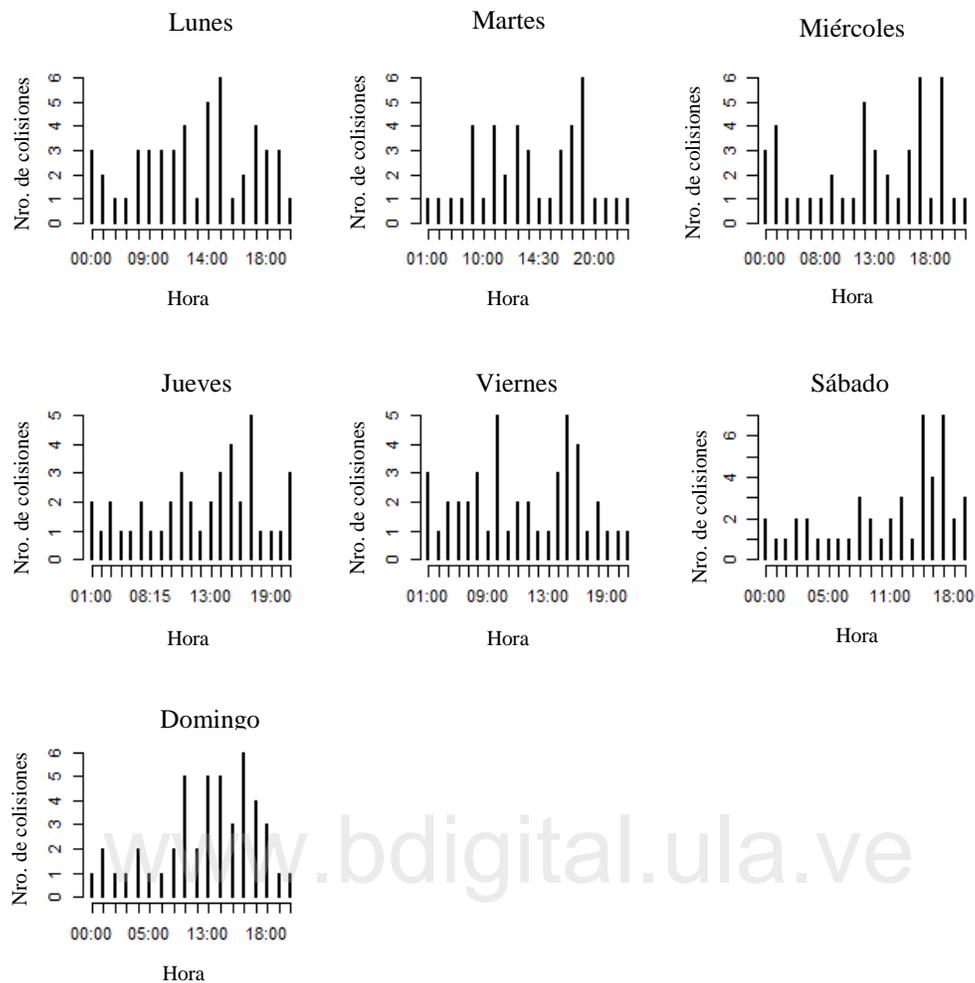


Figura 13. Comportamiento de las horas durante cada día de la semana para el centro Panamericana

A raíz de las pruebas que se han realizado entorno a las colisiones en esta sección, se observa que ahora 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.

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

Tomando los centros que prestan la mayor cantidad de atención a emergencias obtenidos de la prueba de la sección 3.5.1 (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 14).

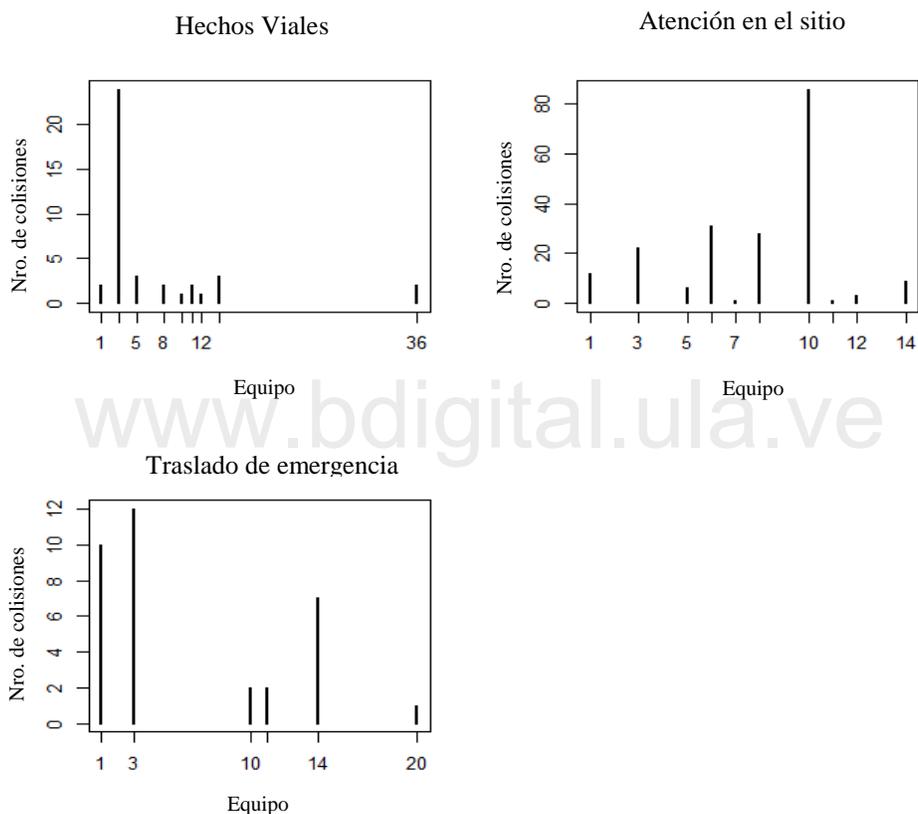


Figura 14. Equipos utilizados para cada actividad no programada en el centro Libertador

En la tabla 17 se presentan los resultados obtenidos, se listan los dos equipos con mayor cantidad de emergencias atendidas, observándose dos puntos importantes:

- Para la actividad atención en el sitio, el equipo con el que se atiende esta actividad principalmente en presencia de colisiones es Sin Unidad para todos los centros bajo estudio.
- Para el centro Libertador se observó un comportamiento diferente al resto; en el caso de los hechos viales se utiliza el equipo SAR seguido de Halcón (código para referirse a la flota de motos de la que dispone INPRADEM). Así mismo, para el traslado de emergencia, se usa más el equipo SAR que el Rojo contrario al comportamiento del resto de los centros.

Tabla 17. 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
Mocotés	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

Estudiados los puntos más críticos de la organización, es decir, los momentos en que se presentan colisiones, aquellas actividades que no pueden ser atendidas inmediatamente, se retoma la prueba de los datos sin incluir las colisiones, pero siempre tomando las actividades no programadas de la organización, las emergencias. Las pruebas relacionadas a colisiones pueden indicar si se necesitan más equipo, más personal, entre otros. La prueba a continuación puede indicar el “cuando”, al estudiar los datos por año.

3.7. 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. Los resultados obtenidos para la semana número 2 y 11 tomando todos los centros durante el período bajo estudio se presentan en las figuras 15 y 16, respectivamente.

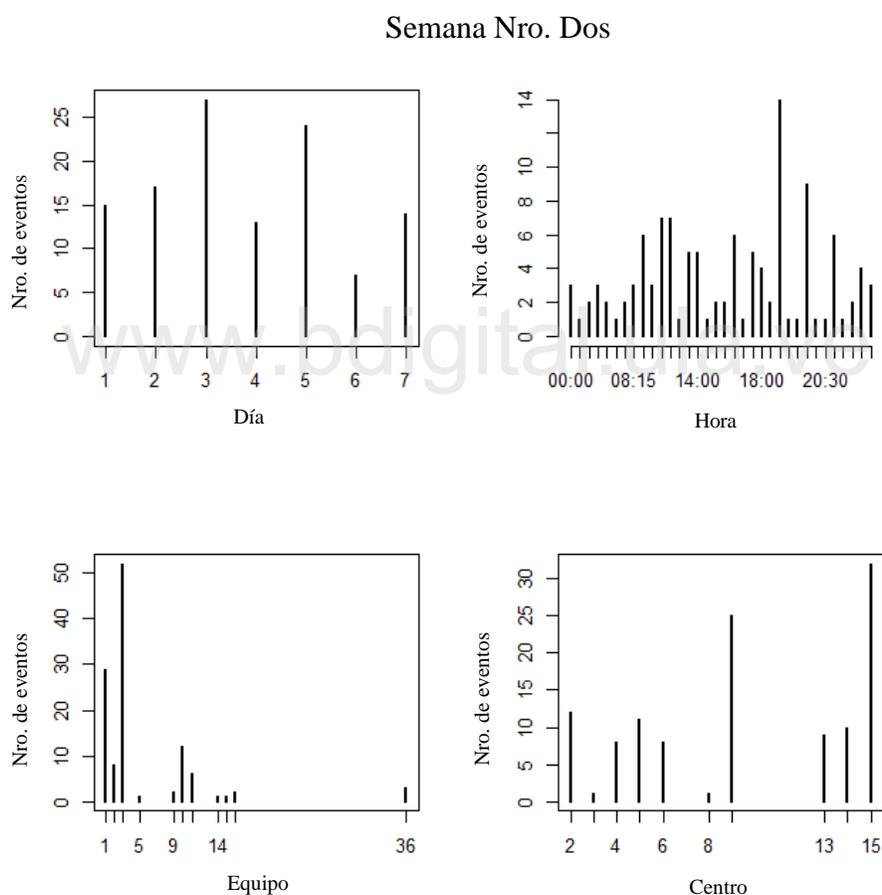


Figura 15. Comportamiento de las variables para la semana número dos del año

Semana Nro. Once

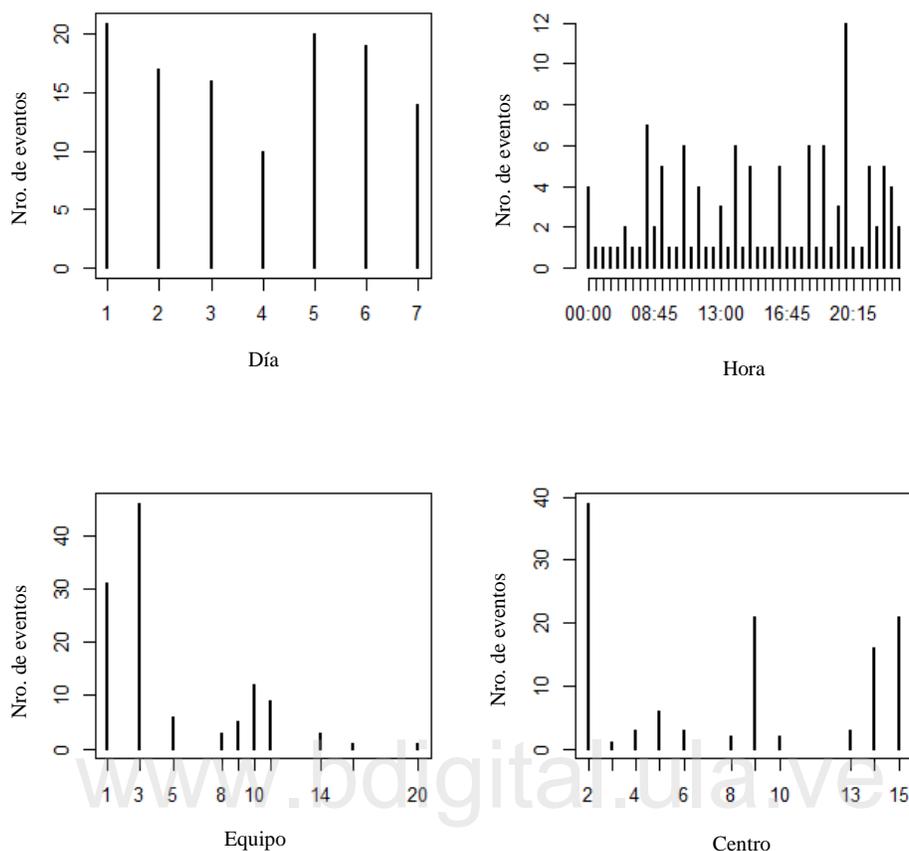


Figura 16. Comportamiento de las variables para la semana número once del año

Se observa que el comportamiento de estas semanas difiere para las variables día y hora, que son claves al momento de crear un plan de gestión que permita atender todas las emergencias de manera óptima. Al mismo tiempo, se descubrió que el equipo que se utiliza en la atención de las emergencias en las semanas donde ocurre más demanda es el mismo, lo cual sugiere enfocarse en las variables día, hora y el centro que poseen mayor demanda; para futuros análisis.

En la tabla 18 se muestra el resumen de las semanas indicando los mayores valores para cada variable. Los valores están ordenados de mayor a menor. La variable día varía para las distintas semanas, aunque las semanas 11, 37 y 38 tienen los mismos días como aquellos en los que se presentan más emergencias. En cuanto a la hora, cuatro semanas presentan un horario nocturno entre siete y ocho y quince de la noche, con excepción de la

semana 43, la cual tiene una y media, y nueve de la mañana como las horas con mayor demanda. Por otra parte, se observa que el equipo utilizado no varía, SAR y Rojo son los más usados para todas las semanas con mayor demanda. Finalmente, Panamericana es el centro que siempre está presente como uno de los más ocupados.

Tabla 18. 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 pm	Hora de Inicio	8:15 pm
Equipo	SAR, Rojo	Equipo	SAR, Rojo
Centro	Paramo, Mocotíes, 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

3.8. 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. Los resultados obtenidos se pueden apreciar en las tablas 19 y 20, respectivamente, y en la figura 17.

Tabla 19. Resultado de las variables para la prueba de las Ferias del Sol

Ferias del Sol	
Variable	Valores
Día	Lunes y Jueves
Equipo	SAR
Hora de Inicio	3:30 pm

Tabla 20. Valores de las variables obtenidos en el prueba de las 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 6:00pm	7:00 y 6:00pm
Equipo	SAR, Rojo	Rojo, SAR	SAR, Rojo
Centro	Panamericana Libertador	Sur del Lago Libertador	Panamericana Sur del Lago
Municipio	Tovar Caracciolo Parra	Tulio Febres Cordero Libertador	Tovar Caracciolo Parra

La tabla 20 resume los resultados obtenidos para las temporadas vacacionales conformados por las pruebas de la época navideña, semana santa y la temporada de agosto. Para la época navideña resalta el día jueves, y un horario tarde noche, durante semana el día lunes y un horario vespertino, y para la temporada de agosto el día sábado en un horario entre los encontrados en Semana Santa y navidad seis y siete de la noche. Para las variables Equipo, Centro y Municipio los valores se repiten para las diferentes épocas, en cuanto a la variable equipo SAR y Rojo son los usados como respuesta a emergencias, en cuanto a centro Panamericana, Libertador y Sur del Lago son los tres centros que presentan mayor demanda en épocas vacacionales y para Municipio Tovar y Caracciolo Parra son el par de municipios en los cuales de presentan más emergencias en época navideña y temporada de agosto, pero en semana santa Tulio Febres Cordero y Libertador son aquellos municipios en donde hay mayor cantidad de emergencias.

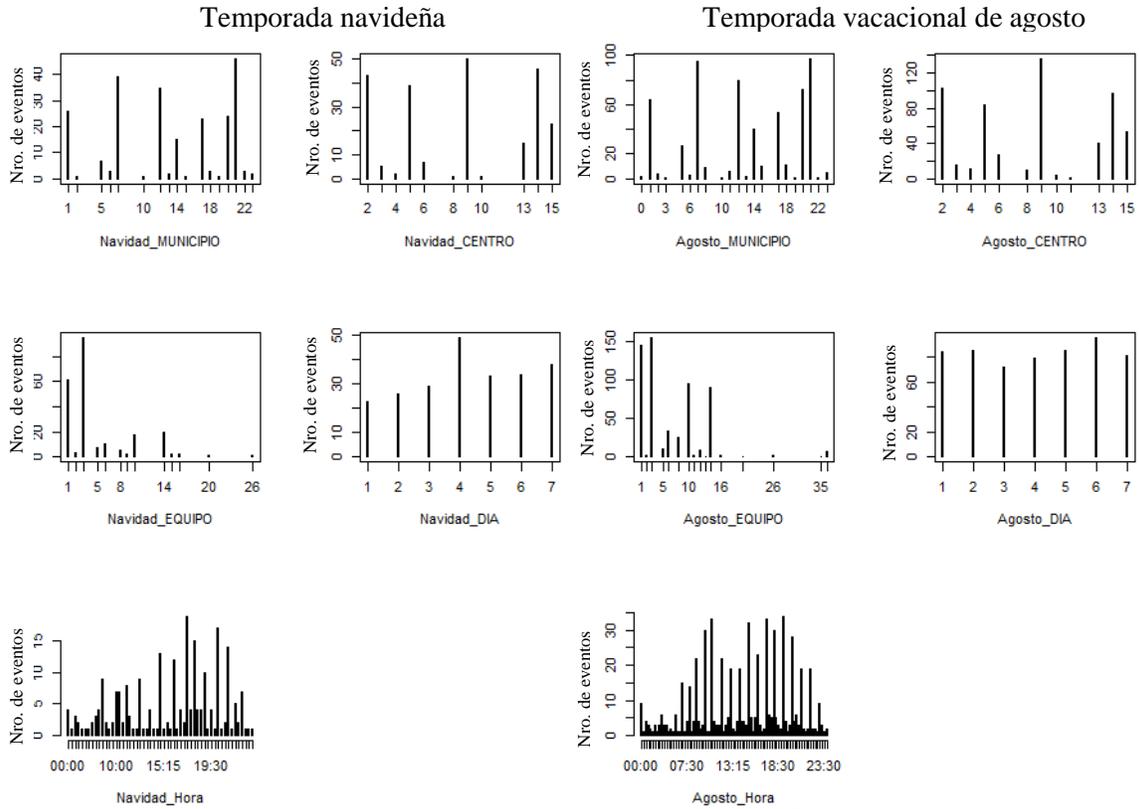


Figura 17. Comportamiento de las variables para la época navideña y la temporada vacacional de agosto

A partir de este estudio se pudo observar que el comportamiento de las horas para las diferentes temporadas vacacionales es diferente, al igual que los días. Por otra parte, se observó que los centros que atienden mayor número de emergencias durante las temporadas de alta demanda son Panamericana, Libertador y Sur del Lago.

Con esta prueba concluye la parte experimental del proyecto, que consistió en la búsqueda de momentos críticos para la organización, es decir, las temporadas en las cuales hay mayor demanda, así como también; el descubrimiento de las horas en las que se presentan un número de emergencias superior al que INPRADEM puede atender. Todo esto, con miras a crear un análisis que permita al final de este proyecto brindar información a la organización que les sea de utilidad para mejorar sus servicios para con la ciudadanía.

Por último, cabe destacar que para cada prueba realizada, solo se muestra una parte de los resultados arrojados. 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

www.bdigital.ula.ve

CAPÍTULO 4. ANÁLISIS DE RESULTADOS

En este capítulo se presenta el análisis de resultados correspondiente a cada uno de los pruebas realizadas. Para ello, se consideran cinco secciones, de acuerdo al orden contemplado en el marco metodológico. La primera sección centrada en el descubrimiento que trazó el camino de las pruebas que siguieron. La segunda, los pruebas de las actividades no programadas. Tercero, las pruebas realizadas en los puntos críticos encontrados, es decir, las colisiones de las horas en la atención de emergencias. Cuarta, el prueba de las semanas que presentan mayor demanda durante el año. Finalmente, una quinta sección enfocada a las temporadas vacacionales.

4.1. Análisis de pruebas iniciales

Las pruebas iniciales de frecuencia para cada variable orientaron la atención hacia la variable horas, observándose con 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 18).

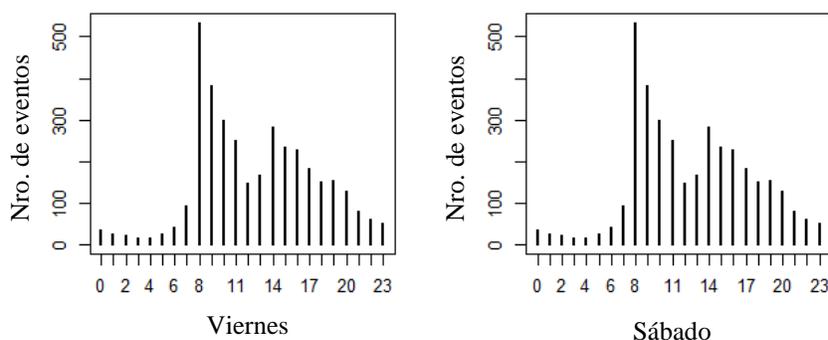


Figura 18. Comportamiento de los días viernes y sábado para la prueba 3.4.2

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, descubriéndose 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, estaba afectando la prueba general.

4.2. Análisis de pruebas entorno 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.

Por otra parte, 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 19).

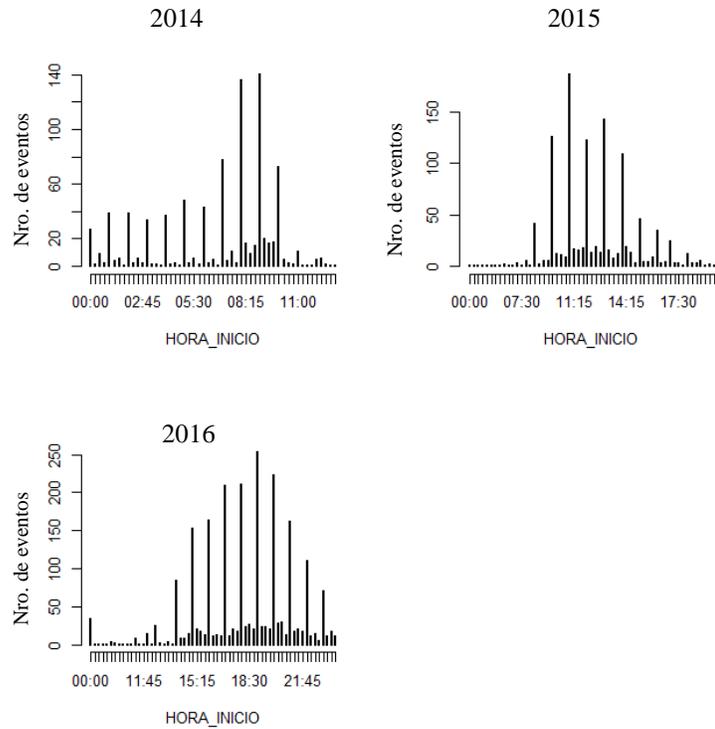


Figura 19. Horas de las emergencias presentadas para cada año

Los resultados anteriores permitieron obtener un esquema de comportamiento general de cada variable (ver Figura 20).

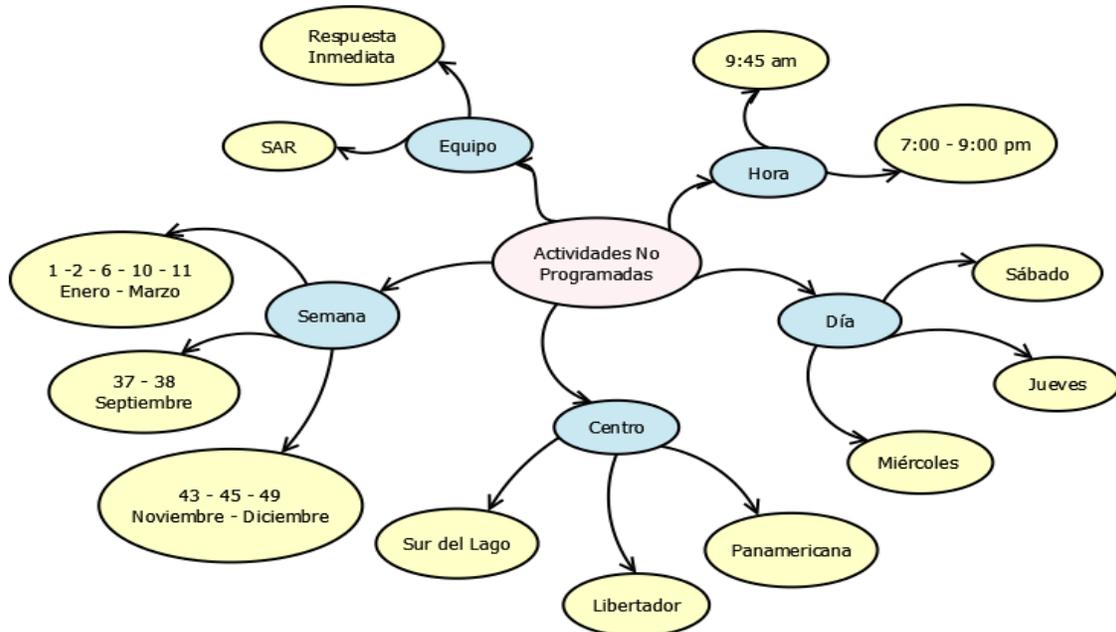


Figura 20. Esquema de comportamiento de las emergencias

Tomando como referencia los resultados obtenidos de la prueba por año se decidió buscar aquellos puntos en donde la hora de inicio es igual a la hora final, considerando un intervalo de tiempo (quince minutos) para considerar que se producen colisiones, observándose que los puntos máximos de la prueba sin colisiones no se reflejan en el prueba de colisiones como era de esperarse. Por tal motivo, se realizaron las pruebas considerando las colisiones.

4.3. Análisis de los pruebas entorno a las colisiones

La prueba de las colisiones en los diferentes centros para las variables Hora y Día (ver Figura 21); 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 (por cada día de la semana se estudiaron las horas en que se presentan las colisiones) considerando aquellos centros que atienden mayor número de emergencias (ver Figura 22) permitió la creación de la tabla 14.

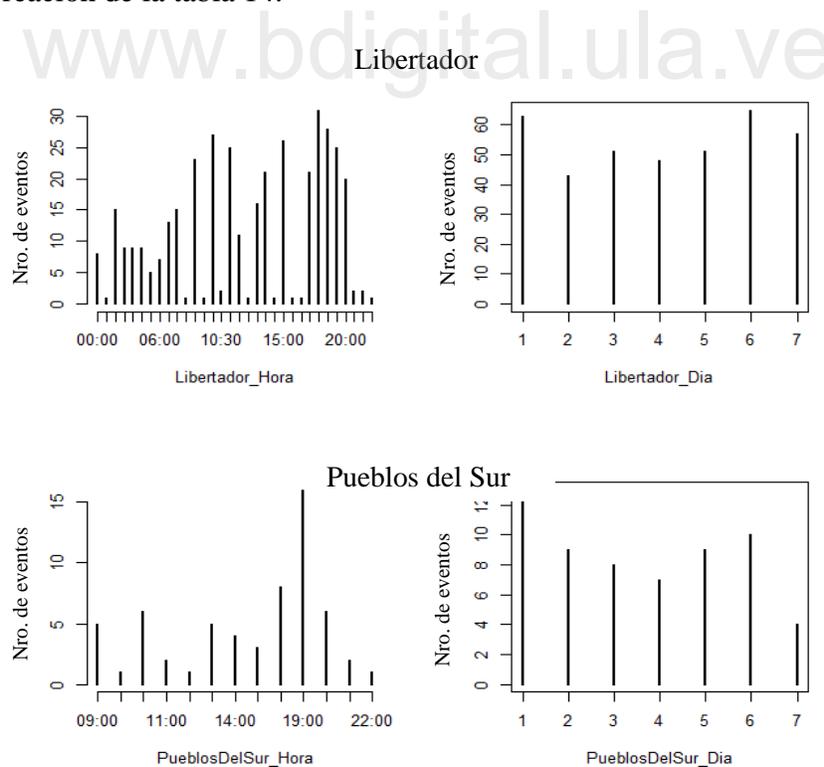
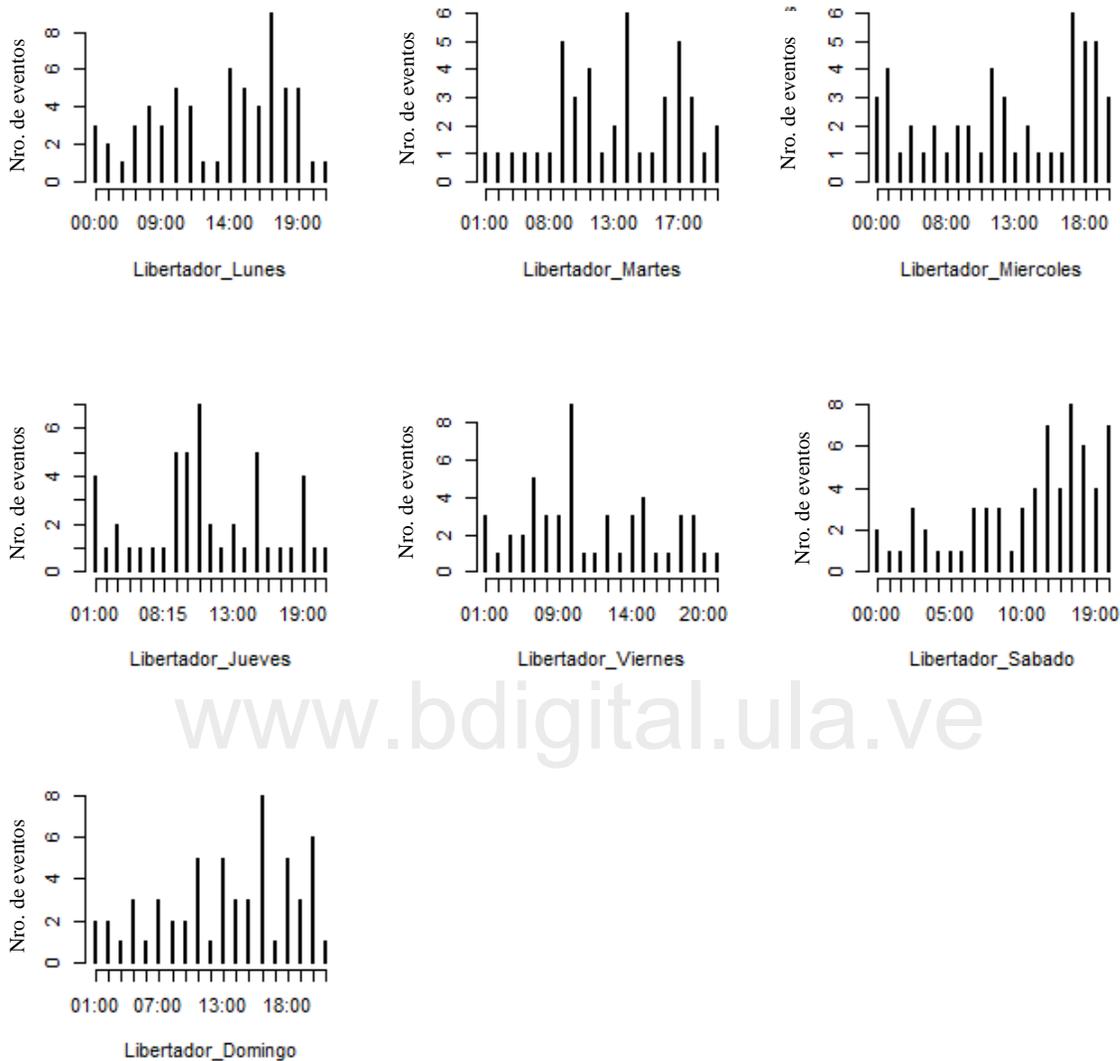


Figura 21. Para los centros Libertador y Pueblos del Sur, comportamiento de las colisiones para las variables Hora y Día de la prueba 3.6.1



www.bdigital.ula.ve

Figura 22. Comportamiento de las horas durante cada día de la semana para el centro Libertador de la prueba 3.6.2

En la tabla 21 se presentan los puntos críticos por día de la semana para el centro Libertador, observándose lo siguiente:

- Durante la semana, en su mayoría existen colisiones en el bloque del llamado “salida del trabajo” o “regreso a casa” entre las seis y siete de la noche. Por otra parte, 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.

Tabla 21. Tabla de colisiones para el centro Libertador

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

En la tabla 22 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 las tarde.
- Un último bloque en la tarde noche, que va desde las cinco hasta las siete de la noche.

Tabla 22. Tabla de colisiones para el centro Panamericana

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

Finalmente, en la tabla 23 se reflejan los resultados para el centro Mocotíes, el cual está encargado de atender las poblaciones de Tovar y Zea. Este centro presenta 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 de la tarde se producen las colisiones. Tiene sentido que esta hora se encuentre reflejada, debido a que este centro se encarga de zonas en donde la jornada laboral culmina más temprano comparado con otros centros como por ejemplo Libertador.

Tabla 23. Tabla de colisiones para el centro Mocotíes

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

Una vez identificados los bloques de horas para los centros, se realizó la prueba 3.6.3, orientada a descubrir los equipos más utilizados en presencia de colisiones con el objeto de que dichos equipos estén disponibles para cubrir la demanda, observándose que los equipos 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. En tal sentido, se observó la necesidad de un plan de gestión que les permita contar con los equipos necesarios para atender todas las emergencias. Adicionalmente, se observó que el centro Libertador utiliza el equipo Halcón (flota de motos) para atender la emergencia hechos viales.

En este punto, otra prueba de interés era analizar cada unidad que conforma los diferentes grupos de equipos, pero no fue posible porque INPRADEM al momento de recolectar los datos no detalla el uso específico de cada unidad, es decir, una ambulancia que pertenece al código Rojo, debería poseer un código propio que permita conocer si esa ambulancia en particular fue usada en la atención de una emergencia, pero no lo tiene.

4.4. Análisis de las semanas con mayor cantidad de emergencias

Para lograr el objetivo de este proyecto de brindar información que de soporte al proceso de toma de decisiones a INPRADEM es menester estudiar los puntos críticos. El primero de ellos, las colisiones en el caso de las horas, las cuales ya fueron analizadas. En segundo lugar, es necesario brindar información para aquellas semanas en las que se presentan mayor cantidad de emergencias durante el año.

En la tabla 24 se presentan los resultados para dicha prueba, 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 observa que Panamericana hace presencia en todas las semanas, aportando una posible relación entre esas horas y el centro.
- La información resumida en dicha tabla 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.

Tabla 24. Comportamiento de las variables para las semanas con más eventos

	Comportamiento de las semanas con más emergencias													
	Hora(24 horas)			Día					EQ	Centro				
	8 - 9	13-15	19-20	L	M	MI	J	V	S	Sar Rojo	Pana.	Mocot.	SurLa.	Páramo
Semana 2														
Semana 11														
Semana 37y 38														
Semana 10														
Semana 43														

4.5. Análisis de las emergencias en temporadas vacacionales

El Estado Mérida con su característica de Estado turístico, anteriormente mencionado, tiene una gran afluencia de personas durante ciertas temporadas a lo largo del año, esto trae como consecuencia el aumento de posibles emergencias. Es por ello, que se decidió hacer una prueba para observar el comportamiento de las variables durante estas fechas y permita a INPRADEM mejorar sus servicios. El análisis para la tabla 25 se presenta a continuación:

- 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.
- Para 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.
- La variable equipo no se expresa en los resultados, como se mencionó previamente, SAR y Rojo son los dos equipos más usados en la atención de emergencias.

Tabla 25. Comportamiento de variables día y hora para las temporadas vacacionales

	Día							Hora (Formato de 24 horas)				
	L	M	MI	J	V	S	D	4-6	8-10	14-16	17-19	20-22
Navidad												
Semana Santa												
Agosto												

Por otro parte, el centro que atiende mayor número de emergencias durante épocas vacacionales es Panamericana, reafirmando su alta demanda, ya que en cada realizado siempre estuvo marcando una gran diferencia con respecto a los demás centros. En segundo lugar aparece el centro Libertador, este comportamiento es el esperado, al ser la ciudad de Mérida de especial atracción turística por el Teleférico, una de las atracciones principales con las que cuenta el Estado (ver Tabla 26).

Tabla 26. Comportamiento de la variable Centro para las temporadas vacacionales

	Panamericana	Libertador	Sur del Lago	Mocotíes
Navidad				
Semana Santa				
Agosto				

El esquema final de puntos críticos para cada temporada vacacional se presenta en las figuras 23, 24 y 25, respectivamente, observándose 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.

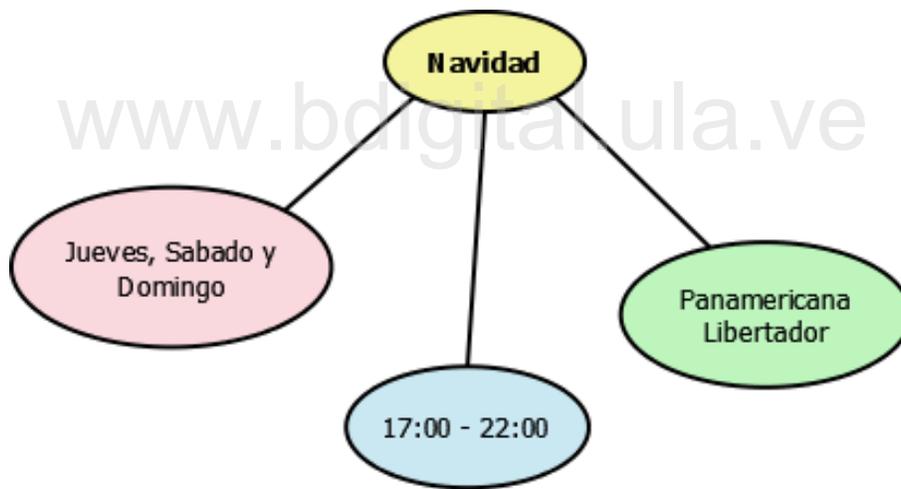


Figura 23. Resumen de resultados para la temporada navideña

- 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.

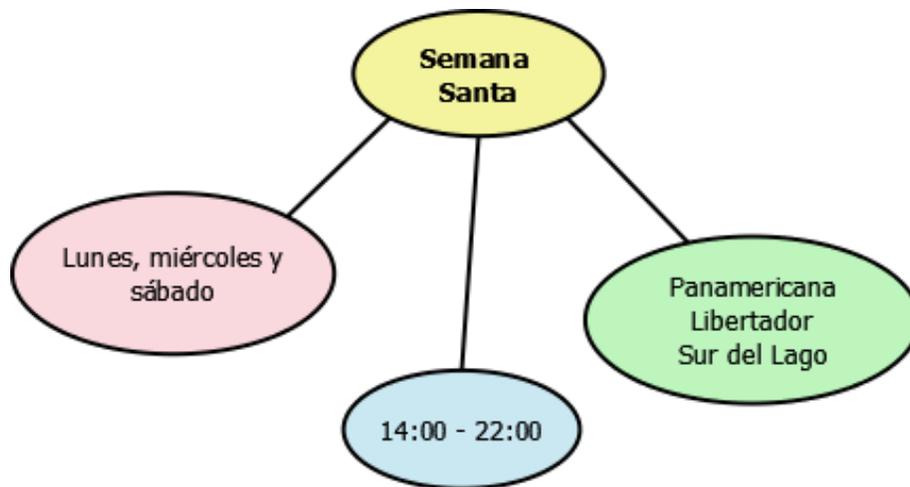


Figura 24. Resumen de resultados para la temporada de Semana Santa

- 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 la que se presentan emergencias se comportan de manera muy parecida a la temporada navideña para la tarde-noche, pero agosto presenta la particularidad, es el único período vacacional que presenta una alta cantidad de emergencias en horas de la mañana.

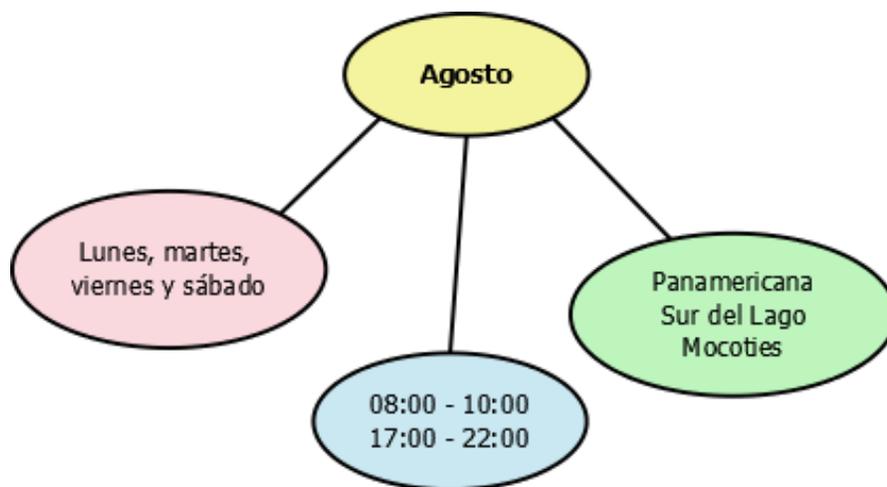


Figura 25. Resumen de resultados para la temporada vacacional de agosto

CAPÍTULO V. CONCLUSIONES Y RECOMENDACIONES

En este último capítulo se presentan las conclusiones una vez realizadas las pruebas experimentales y el análisis de las mismas. Además, se exponen algunas recomendaciones para trabajos futuros.

5.1. Conclusiones

Este capítulo equivale a la fase 5 y 6 de la metodología empleada en este proyecto, evaluación y despliegues de resultados, corresponde a la evaluación del estudio experimental realizado con respecto a los objetivos del proyecto, así como también se evalúa su aplicación.

Los objetivos planteados al comienzo del proyecto se cumplieron satisfactoriamente. Se logró caracterizar los registros de actividades de INPRADEM, profundizando los conocimientos con respecto al comportamiento de las actividades que se realizan en la organización. El conocimiento es presentado por medio de una serie de tablas y figuras que permiten representar el comportamiento de las variables a través de los diversas pruebas realizadas.

Se inició con el estudio de los formatos de recolección de información, en este caso se refiere a la planilla en físico y digital que utilizan en la organización, de modo que se determinara la mejor manera de prepararlos, y se analizaron para asegurarse si se podría realizar una prueba de caracterización utilizando minería de datos.

Seguidamente, se modificó la presentación de los datos originales para que fuesen utilizables en RStudio, que fue la herramienta usada para el desarrollo de este proyecto. Los datos fueron transformados a números, asignándole un código a cada valor nominal de las variables. Con la base de datos ya preparada, se definieron las estrategias de caracterización de los datos, éstos por ser de carácter nominal dirigieron el proyecto a trabajar con gráficos de frecuencia, siendo la prueba más recomendada para el estudio de

descripción en minería de datos, específicamente en este trabajo la caracterización de variables nominales.

Se implementaron según el objetivo de este proyecto para el análisis de datos, los algoritmos utilizados en minería de datos llamados no supervisados o de descubrimiento del conocimiento, buscando encontrar patrones o tendencias que lleven 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 los 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.

Se realizaron pruebas en torno a las emergencias para lograr encontrar más información sobre estas actividades en particular, enfocando el objetivo de la prueba en aquellas horas donde es probable que la emergencia no haya sido atendida inmediatamente, a estos casos se les llamo colisiones. Se llevó a cabo una prueba de las emergencias centrada en encontrar comportamientos de las variables durante las colisiones. En estos resultados algunas variables al ser analizadas brindaron valiosa información: se encontraron 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; así mismo se encontró que el equipo SAR y Rojo son los más usados en la atención de emergencias, pero durante las colisiones en algunos casos se atendían las emergencias sin equipo alguno y solo son despachados oficiales de atención de emergencias; a partir de la cantidad de colisiones que se presentan para un mismo equipo se concluyó la necesidad de una gestión que disponga de mejor manera los equipos y personal.

Se realizó un amplio estudio entorno a las semanas en las cuales se presentan la mayor cantidad de emergencias, como de las semanas que conforman las temporadas vacacionales que se desarrollan a lo largo del año. Estas son pruebas necesarias para mejorar la gestión de la organización especialmente durante estas semanas donde se esperan aumento de emergencias. Para estas semanas en particular se realizaron varias pruebas obteniendo como resultado 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. La prueba de las variables tomando como partida la

variable semana para algunos valores en específico, permitió un análisis de resultados que llevaron a la creación de tablas y esquemas en donde se aprecia el comportamiento de diferentes variables, 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, así mismo 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., de modo que puedan concentrarse en estos resultados como prioridad al momento de crear dicho plan.

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.

Recordemos 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.

5.2. Recomendaciones

- INPRADEM tiene la disposición de adentrarse en el mundo de la obtención de información por medio de recolección de datos. La prueba de ello fue la creación de la planilla de los registros de las actividades realizadas en la organización. Para realizar mejores pruebas, que puedan otorgar información a más detalle y descubrir patrones e incluso realizar predicciones se recomienda la creación de una planilla diseñada por un experto, que permita una extracción de información de mayor calidad.
- Se recomienda a la organización otorgar a cada funcionario que se desempeña en las actividades que se desarrollan, un código específico que permita realizar pruebas desde el punto de vista de recursos humanos.
- Así mismo, asignar código a cada unidad que forma parte del código general de la variable equipo, 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 en momentos en los cuales se necesiten en otras ubicaciones.

- Realizar un proyecto que se encargue de crear y desarrollar un plan de gestión en conjunto con la organización que permita optimizar el tiempo y la gestión de los equipos utilizados en emergencias.
- Desarrollar una prueba entorno a las actividades de planeación previa en conjunto con INPRADEM que sirva de soporte para la distribución óptima de este servicio, para mayor comodidad de la organización y de la ciudadanía que demanda el servicio.
- Desarrollo de una aplicación web que permita que la recolección de los datos sea netamente digital, pues existe la posibilidad de perder datos en la transcripción.

www.bdigital.ula.ve

REFERENCIAS

Aguilera, E. (2009). *Seguridad vial en Venezuela*. Libro creado, desarrollado y editado en Venezuela por Fundación Seguros Caracas © 2009. Primera edición. Caracas, Venezuela. Disponible en http://www.intt.gob.ve/repositorio/biblioteca/educacion_y_seguridad_vial/Seguridad_Vial_Prof_Elio_Aguilera.pdf

Ahlemeyer-Stubbe, A. y Coleman, S. *A Practical Guide to Data Mining for Business and Industry*. Editorial Wiley. Año 2014.

Altamiranda, L., Peña, A. y Ospino De La Rosa, M. (2013). *Minería de datos como herramienta para el desarrollo de estrategias de mercadeo*. Universidad Autónoma del Caribe, Barranquilla, Colombia. Disponible en <http://revistas.uexternado.edu.co/index.php/sotavento/article/view/3709/3978>

Álvarez, A y Chica, S. (2008). *Gestión de las organizaciones públicas*. Escuela Superior de administración pública. Bogotá, Colombia.

Aránguiz, A. (2012). *Análisis de accidentes de tránsito en zonas urbanas y rurales usando minería de datos difusa*. Pontificia Universidad Católica de Valparaíso, Facultad de Ingeniería. Valparaíso, Chile.

Bonilla, H., Jaramillo (2009). Llanos,R. *Research in social sciences*. Universidad del Norte. Colombia

Britos, P. (2008). *Procesos de explotación de información basados en sistemas inteligentes*. Universidad Nacional de La Plata. Argentina. Disponible en http://sedici.unlp.edu.ar/bitstream/handle/10915/4142/Documento_completo.pdf?sequence=1.

Carrillo, J. Virseda, F. *Mineria de datos y aplicaciones*, Departamento de ingeniería telemática. Universidad Carlo III de Madrid. 2007.

Chapman, P, Clinton, J. Khabaza, T. Reinartz, T. Rüdiger, W. *The CRISP-DM Process*. 2000.

Cichosz, P. *Data mining Algorithms: Explained Using R*. Department of Electronics and Information Technology Warsaw University of Technology. Poland. Editorial Wiley. 2015.

De La Vara Salazar, R. y Gutierrez H. *Análisis y diseño de experimentos*. Editorial McDraw-Hill. Segunda edición.

Frawley, W. *Knowledge Discovery in Databases: An Overview*. AI Magazine Volume 13, Number 3. 1992.

García, R. A., Delgado, D., Díaz, E. y García, R. R., (2012). *Caracterización de la accidentalidad vehicular y análisis de las causas en la provincia de Villa Clara, Cuba*. Disponible en <http://revistas.unal.edu.co/index.php/dyna/article/view/20038/43525>.

Gordillo, S. y Haedo A. (2011). *Estudio comparativo de metodologías para minería de datos*. Facultad de Ciencias Exactas de la Universidad Nacional de Buenos Aires y Facultad de Informática de la Universidad Nacional de La Plata. Disponible en <https://digital.cic.gba.gob.ar/bitstream/handle/11746/3525/Estudio%20comparativo%20de%20metodolog%C3%ADas%20para%20miner%C3%ADa%20de%20datos.pdf-PDFA.pdf?sequence=1>.

Han, J., Kamber, M. (2001). *Scalable frequent-pattern mining methods: an overview*. Simon Fraser University. Canada.

Hassinger, (2015). *Aplicación de técnicas de minería de datos en accidentes de tráfico*. Universidad Politécnica de Valencia, Departamento de sistemas informáticos y computación. Valencia, España.

Instituto nacional de estadística (INE). *Cifras del parque automotor del estado Mérida*. Disponible en <http://www.ine.gov.ve/>.

Koontz, H., Weihrich, H. *Administración, una perspectiva global*. Editorial Mc Graw Hill. México, 1999.

León, E. (2017). *Módulo Minería de Datos*. Universidad Nacional de Colombia. Grupo de investigación MIDAS. Disponible en http://disi.unal.edu.co/~eleonguz/cursos/md/presentaciones/Sesion5_Metodologias.pdf

Logreira, C. *Minería de datos y su incidencia en la toma de decisiones empresariales en el contexto de CRM*. Revista Ingeniería solidaria Volumen 7, números 12-13. Enero 2011.

Marcano, Y. y Talavera R. (2007). *Minería de Datos como soporte a la toma de decisiones empresariales*. Universidad del Zulia, Núcleo Punto Fijo. Disponible en http://www.scielo.org.ve/scielo.php?script=sci_arttext&pid=S1012-15872007000100008

Martínez, G. y Cintas del Río, R. (2015). *Metodología de minerías de datos para el estudio de tablas de siniestralidad vial*. Universidad Complutense Madrid, España. Facultad de estudios estadísticos.

Moine, J., Haedo, A y Gordillo, S. *Estudio comparativo de metodologías para minería de datos*. Facultad de Informática, Universidad Nacional de La Plata. Argentina.

Montgomery, D. *Diseno y análisis de experimentos*. Universidad estatal de Arizona. Editorial Limusa Wiley. Segunda edición.

Protección Civil y Administración de Desastres del Estado Mérida. *Reseña histórica, misión, visión, objetivos de INPRADEM*.

Riquelme, J., Ruiz, R. y Gilbert, K. (2006). *Minería de Datos: Conceptos y Tendencias*. España. Red de Revistas Científicas de América Latina y el Caribe, España y Portugal. Disponible en <http://www.redalyc.org/html/925/92502902/index.html>.

Salinas, M. y Vele, L. (2014). *Estudio científico de la accidentalidad de tránsito en El Cantón Cuenca*. Universidad Politécnica Salesiana, sede Cuenca. Ingeniería Mecánica automotriz. Cuenca, Ecuador.

Sánchez, A. Revista Virtual Universidad Católica del Norte. Número 29. 2010. *El artículo sistematización de experiencias: construcción de sentido desde una perspectiva crítica*.

UNAM Facultad de Contaduría, Administración e Informática. Ciudad de México. *Minería de Datos*. Disponible en https://es.slideshare.net/bemagualli?utm_campaign=profiletracking&utm_medium=sssite&utm_source=ssslideview.

Uzcátegui, E. *Organizaciones públicas*. Encuentro de saberes para el conocimiento. Artículo 31. Año 2009. Colombia.

Wright y Ashill, *Knowledge discovery in databases: tools and techniques*, Publicado en la revista Crossroads, Volumen 5, Issue 2, Diciembre 1998, Nueva York, Estados Unidos.