

DOCUMENTACIÓN TÉCNICA ANALÍTICA DE DATOS METODOLOGÍA CRISP – DM

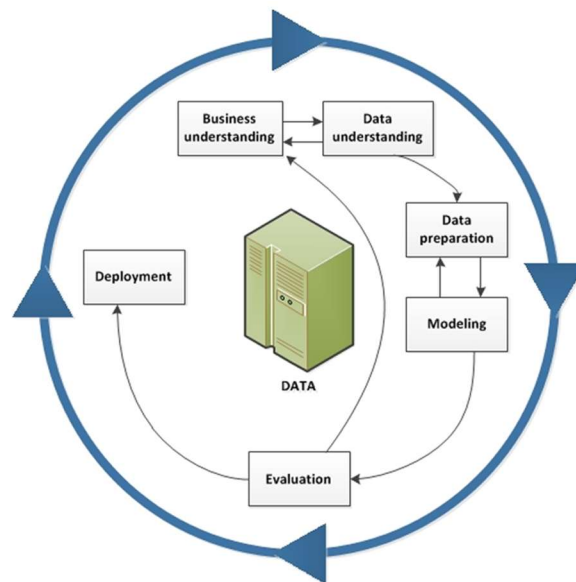
ALCALDIA DE CHIA

2024 - 2025

DESCRIPCIÓN DE LA METODOLOGÍA

La metodología seleccionada por parte de la Administración Municipal, es CRISP-DM (Cross Industry Standard Process for Data Mining), el cual está orientado como metodología dado en fases de proyecto para el desarrollo de acciones de comprensión, modelado y evaluación de datos como respuesta a las necesidades de tomas de decisiones y producción.

La metodología incluye las siguientes fases:



Tomado de: Guía de CRISP-DM de IBM SPSS Modeler. © Copyright International Business Machines Corporation

De acuerdo a lo anterior, el presente documento, contará con la documentación técnica de evolución a través del tiempo de los datos en flujo de dos proyectos de minería de datos comprendidos como: Observatorio de Seguridad y Observatorio de Movilidad, del municipio de Chía.

1. COMPRENSIÓN DEL NEGOCIO

En ésta sección se deben elaborar textos asociados a:

Tarea 1: Determinar la estructura de la organización (Mision y visión en texto, objetivos estratégicos de la alcaldía, Decreto 40 respecto a la Dirección de Seguridad y la Secretaría de Movilidad)

La Alcaldía de Chía es la entidad administrativa principal del municipio, encargada de dirigir y ejecutar las políticas públicas, gestionando los recursos locales para el bienestar de la población. Es la máxima autoridad ejecutiva a nivel municipal, responsable de la administración de la ciudad y de garantizar el cumplimiento de las leyes y normativas aplicables, su misión Garantizar a sus habitantes una oportuna y efectiva prestación de los servicios con calidad en materia de salud, educación, seguridad, construcción de obras de infraestructura, ordenamiento territorial, medio ambiente, crecimiento socio -cultural, deportivo y erradicación de la pobreza, promoviendo la participación comunitaria en aras de mejorar la calidad de vida de nuestros ciudadanos y de quienes visitan nuestro territorio, su visión en el año 2027 Chía será ciudad líder en desarrollo sostenible caracterizada por una cultura emprendedora, empoderada del medio ambiente, participativa, solidaria y orgullosa de su patrimonio e historia.

Los principales objetivos de la Alcaldía de Chía son:

Mejorar la calidad de vida de los habitantes:

Esto implica la atención a áreas como salud, educación, seguridad, infraestructura y servicios públicos.

Gestionar eficientemente los recursos públicos:

Esto incluye la transparencia en la administración, la planificación estratégica y la optimización de la inversión.

Promover la participación ciudadana:

Esto implica fortalecer los espacios de diálogo, concertación y participación de los ciudadanos en la toma de decisiones.

Fortalecer la identidad cultural y patrimonial:

Esto implica la preservación y promoción del patrimonio histórico, cultural y natural de Chía.

Impulsar la transformación digital y la innovación:

Esto implica el uso de la tecnología para mejorar la eficiencia de los servicios públicos y la calidad de vida de los ciudadanos.

Impulsar el desarrollo económico y la creación de empleo:

Esto implica la atracción de inversiones, la promoción de la innovación y el emprendimiento

ESTRUCTURA ORGÁNICA DEL MUNICIPIO

ARTÍCULO 5°. - DE LA ESTRUCTURA ORGÁNICA DEL MUNICIPIO: La estructura Administrativa y funcional del Municipio de Chía está integrada por la siguiente estructura administrativa:

ESTRUCTURA INTERNA DE LA ALCALDÍA MUNICIPAL DE CHÍA

1. DESPACHO DEL ALCALDE

- 1.1. Oficina Asesora Jurídica
- 1.2. Oficina de Contratación
- 1.3. Oficina de Defensa Judicial
- 1.4. Oficina de Control Interno
- 1.5. Oficina de Tecnologías de Información y las Comunicaciones, TIC
- 1.6. Oficina Asesora de Comunicación, Prensa y Protocolo

2. SECRETARÍA DE PLANEACIÓN

- 2.1 Dirección de Sistemas de Información y Estadística.
- 2.2 Dirección de Planificación del Desarrollo
- 2.3 Dirección de Ordenamiento Territorial y Plusvalía
- 2.4 Dirección de Urbanismo
- 2.5 Dirección de Servicios

3. SECRETARÍA GENERAL

- 3.1 Dirección de Función Pública
- 3.2 Dirección de Servicios Administrativos
- 3.3 Dirección Centro de Atención al Ciudadano
- 3.4 Dirección de Control Interno Disciplinario

4. SECRETARÍA DE GOBIERNO

- 4.1. Dirección de Seguridad y Convivencia Ciudadana
- 4.2. Dirección de Asuntos Étnicos, Raciales, Religiosos y Posconflicto
- 4.3. Dirección de Derechos y Resolución de Conflictos

- 4.3.1 Comisarías de Familia
- 4.3.2 Inspecciones de Policía

5. SECRETARÍA DE HACIENDA

- 5.1 Dirección de Rentas
- 5.2 Dirección Financiera
- 5.3 Dirección de Tesorería

6. SECRETARÍA DE OBRAS PÚBLICAS

- 6.1 Dirección de Infraestructura
- 6.2 Dirección de Programación, Estudios y Diseños
- 6.3 Dirección de Valorización

7. SECRETARÍA DE DESARROLLO SOCIAL

- 7.1 Dirección de Ciudadanía Juvenil
- 7.2 Dirección de Acción Social
- 7.3 Dirección de Cultura

8. SECRETARÍA DE EDUCACION

- 8.1 Dirección de Inspección y Vigilancia
- 8.2 Dirección de Gestión y Fomento a la Educación
- 8.3 Dirección Administrativa y Financiera

9. SECRETARÍA DE SALUD

- 9.1 Dirección de Salud Pública
- 9.2 Dirección de Vigilancia y Control

10. SECRETARÍA PARA EL DESARROLLO ECONOMICO

- 10.1 Dirección de Desarrollo Agropecuario y Empresarial
- 10.2 Dirección de Turismo

11. SECRETARÍA DE MEDIO AMBIENTE

12. SECRETARÍA DE MOVILIDAD

- 12.1 Dirección de Servicios de Movilidad y Gestión del Transporte
- 12.2 Dirección de Educación, Seguridad Vial y Control del Transito

13. SECRETARÍA DE PARTICIPACION CIUDADANA Y ACCION COMUNITARIA

14. ORGANOS DE ASESORIA Y COORDINACION

**ARTÍCULO 39°-. Dirección de Seguridad y Convivencia Ciudadana.
Son funciones de la Dirección de Seguridad y Convivencia Ciudadana:**

1. Apoyar a la Secretaría en el trazado de políticas de prevención de delitos, contravenciones y problemas de convivencia y seguridad ciudadana.
2. Adelantar los estudios conducentes para la formulación de los planes, programas y proyectos específicos para atender el fenómeno de los desplazados por la violencia en el Municipio y otros fenómenos que afecten a la población civil, gestionando recursos con organismos nacionales e internacionales.
3. Establecer programas orientados a fomentar la solución democrática y pacífica de los conflictos y facilitar la participación de todos en las decisiones que los afectan.
4. Promover en la administración municipal la resolución de manera pronta las solicitudes que los ciudadanos les presenten para la satisfacción de sus necesidades y la prevención y tratamiento de las perturbaciones a la seguridad, la tranquilidad, la salubridad y el ambiente.
5. Diseñar y formular proyectos y programas de prevención de delitos, contravenciones y problemas de convivencia y seguridad ciudadana.
6. Organizar el observatorio de convivencia y seguridad.
7. Estimular la participación de la ciudadanía en la adopción de mecanismos de colaboración con las autoridades de policía y justicia, en especial para la reducción de la delincuencia y el consumo de las sustancias psicoactivas.
8. Apoyar al Secretario de Gobierno en el diseño y desarrollo de los planes y estrategias integrales de seguridad con la Policía Nacional, atendiendo las necesidades y circunstancias de las comunidades en el municipio.
9. Determinar las estrategias relacionadas con las políticas, planes, programas y proyectos relativos al mantenimiento y restablecimiento del orden público y coordinar con las autoridades locales, militares, de policía y los organismos de seguridad la ejecución de las mismas.

10. Coordinar y colaborar, bajo la dirección de las autoridades competentes, los procesos electorales para que se realicen en el municipio, bajo los parámetros de transparencia, armonía, tranquilidad y seguridad.

11. Participar con la Policía Nacional en la identificación de líneas generales de diseño, formulación, seguimiento, evaluación y control de proyectos y programas genéricos o específicos de prevención y control en seguridad ciudadana que deban ser ejecutados por ella, siguiendo las pautas que para el efecto señale el Secretario de Gobierno.

12. Asumir y coordinar la consecución de información de seguridad ciudadana cualitativa y cuantitativa, incluyendo aquella relativa a las reacciones, posturas, propuestas y acciones de otras autoridades y de la sociedad civil.

13. Evaluar y revisar periódicamente el impacto, la pertinencia y la oportunidad de las políticas y estrategias de seguridad ciudadana trazadas por el Gobierno Municipal.

DE LA SECRETARÍA DE MOVILIDAD ARTÍCULO 80°. DE LA MISIÓN Y OBJETIVOS DE LA SECRETARÍA DE MOVILIDAD:

La misión de la Secretaría de Movilidad es generar una movilidad sustentable en armonía con sus características urbanas y rurales, que potencie el desarrollo y la competitividad del municipio, con respeto a los usuarios de las vías: desarrollando condiciones de seguridad, eficiencia, eficacia y economía para la comunidad Chiense con servidores públicos con espíritu de servicio, profesionales y comprometidos.

ARTÍCULO 81°. Organización interna de la Secretaría de Movilidad.

La organización interna de la Secretaría de Movilidad, es la siguiente:

12. SECRETARÍA DE MOVILIDAD

12.1 Dirección de Servicios de Movilidad y Gestión del Transporte

12.2 Dirección de Educación, Seguridad Vial y Control del Tránsito

Tarea 2: Describir el área problemática (Como es la problemática de la captura de datos, o de análisis cruzados etc.)

Desde la problemática de la consecución de datos evidenciamos que con respecto a la base de datos del IPAT implementada por la secretaria de movilidad reporta una recopilación de la información manera manual, donde tanto el funcionario (Agente

de tránsito) como el funcionario encargado de transcribir la información reportado podrían omitir, modificar o dar por perdidos datos fundamentales, desde ese punto de vista se genera como problemática la fiabilidad de los datos recopilados, por otro lado, para la oficina de Defensa Judicial se implementó la plataforma de Participa en Chía, en el aplicativo de recepción de llamadas de emergencia del 123, donde se generan los reportes directamente desde el aplicativo descargándolo en formatos .xlsx, lo cual representa datos totalmente reales y confiables.

Tarea 3: Evaluación de la situación (Como se hace actualmente los dos procesos)

En este momento para la secretaria de movilidad se tiene una recolección de datos de forma manual, donde el Agente de tránsito que conoce el caso relaciona una serie de datos tanto de las personas como de los vehículos involucrados, una vez terminado el caso el Agente radica un reporte con toda la información recopilada en el formato IPAT (informe policial de accidente de tránsito) el cual se encuentra reglamentado en la resolución 11268 de 2012 del Ministerio de transporte, el funcionario de la secretaria de movilidad encargado de recibir la información, transcribe en una base de datos establecida de manera manual todos los datos relacionados los cuales se almacenan en forma local.

Por otro lado, para la Dirección de Seguridad y Convivencia Ciudadana se implemento el aplicativo de recepción de llamadas de emergencia 123, para el cual se capacitó a los funcionarios de la Alcaldía, entes de emergencia (Policía, Bomberos y Transito) y grupo de apoyo ciudadano Fenix para su correcto manejo, en el cual el aplicativo consta de un formulario donde se relaciona toda la información de la llamada, almacenándose de forma automática y donde por ningún motivo se puede alterar la información registrada, a su vez el aplicativo permite descargar informe tanto de una llamada en específico como de diferentes periodos, dichos reportes se generan en formato pdf, lo que se convierte en una fuente de información veraz e inalterable.

1. COMPRENSIÓN DE LOS DATOS

Descripción de los orígenes de datos:

Para la Dirección de Seguridad y Convivencia Ciudadana, la fuente de origen de los datos proviene del aplicativo de recepción de llamadas de emergencia 123, implementado en la plataforma "Participa en Chía". Este sistema permite registrar de manera automática y en tiempo real las llamadas de emergencia, garantizando la integridad y confiabilidad de la información recopilada.

Para la secretaria de movilidad se tiene una recolección de datos de forma manual, donde se relaciona una serie de datos tanto de las personas como de los vehículos involucrados, una vez terminado el caso el Agente radica un reporte con toda la información recopilada en el formato IPAT (informe policial de accidente de tránsito), para que el funcionario de la secretaria de movilidad encargado de recibir la información, transcribe en una base de datos establecida de manera manual todos los datos relacionados los cuales se almacenan en forma local.

Fichas de descripción de los datos:

A continuación, se presentan las fichas de descripción de los datos correspondientes a cada fuente:

Sistema de Recepción de Llamadas de Emergencia 123 - Dirección de Seguridad y Convivencia Ciudadana.

Origen de los datos: Llamadas registradas a través de la plataforma "Participa en Chía".

Formato: Digital, con almacenamiento automático en la base de datos del sistema.

Contenido: Datos de la llamada, incluyendo tipo de emergencia, ubicación, hora, duración, y respuesta de los entes de emergencia.

Proceso de captura: Los operadores ingresan la información directamente en el sistema durante o inmediatamente después de la llamada.

Frecuencia de actualización: En tiempo real, con cada nueva llamada recibida.

Ventajas: Alta confiabilidad y precisión en los datos, con mínima intervención humana.

IDENTIFICACIÓN DE DATOS Fuentes de Extracción	Código: 2501
	Versión: V2
	Página: 1
	Vigente: Si

1. IDENTIFICACIÓN DE LOS DATOS

1.1 Nombre de la Fuente	BD_123.xlsx						
1.2 Descripción detallada de la fuente datos (Detalle a que hacen referencia los datos)	Contempla la identificación y ubicación de las zonas y las regiones georreferenciadas dentro de aplicativo de llamadas de la central de emergencias 123.						
1.3 Tipo de Origen de los Datos	WEB		APLICACION	X	BD		OTRO
1.4 URL o Referencia de origen							
1.5 Cantidad de Variables							
1.6 Cantidad de Registros	5						
1.5 Peso físico – Si aplica	14 Kb						

2. DICCIONARIO DE DATOS

Nombre Campo	Tipo de Dato	Nulo (S/N)	Transformación (S/N)	Descripción
id_zona	Numérico	No	No	Digito de identificación que relaciona una zona en específico proporcionada por el aplicativo.

zona	Texto	No	No	Nombre de la zona determinada dentro del Municipio proporcionada por el aplicativo.
id_tipo_zona	Numérico	No	No	Digito de identificación que relaciona el tipo de zona proporcionada por el aplicativo, ya sea rural, urbano o no definido.
id_region	Numérico	No	No	Digito de identificación que relaciona la ubicación de la región definida en el aplicativo, ya sea oriental, occidental, entre ríos sur, entre ríos norte o no definida.
región	Texto	No	No	Nombre de la región definida en el aplicativo, ya sea oriental, occidental, entre ríos sur, entre ríos norte o no definida.

Llave primaria (Si aplica)	
Llave(s) Foránea(s)	
Otras Restricciones	

IDENTIFICACIÓN DE DATOS Fuentes de Extracción	Código: 2501
	Versión: V2
	Página: 1
	Vigente: Si

1. IDENTIFICACIÓN DE LOS DATOS

1.1 Nombre de la Fuente	Llamadas.xlsx						
1.2 Descripción detallada de la fuente datos (Detalle a que hacen referencia los datos)	Base de datos en archivo plano generada desde la aplicación de recepción de llamadas de la línea de emergencia 123 del Municipio de Chía.						
1.3 Tipo de Origen de los Datos	WEB		APLICACION		BD	X	OTRO
1.4 URL o Referencia de origen							
1.5 Cantidad de Variables							
1.6 Cantidad de Registros	24						
1.5 Peso físico – Si aplica	6.6 Kb						

2. DICCIONARIO DE DATOS

Nombre Campo	Tipo de Dato	Nulo (S/N)	Transformación (S/N)	Descripción
Numcaso	Numérico	No	No	Consecutivo generado por el sistema con el fin de identificar el número del caso.
Usuariorol	Numérico	No	No	Número que identifica el rol de la persona que realiza la llamada.
Fecha	Numérico	No	No	Día, mes y año correspondiente a la recepción del caso.

Nombre	Texto	Si	No	Nombre proporcionado por el usuario al momento de registrar la llamada.
Teléfono	Numérico	Si	No	Número telefónico del usuario que realiza la llamada.
Dirección	Alfanumérico	No	No	Dirección aportada por el usuario del lugar de los hechos.
Tipzona	Numérico	No	No	Número que identifica la zona ya sea rural o urbano.
Región	Numérico	No	No	Número que identifica la región relacionada a los puntos cardinales.
Zona	Numérico	No	No	Número que representa la zona y relaciona a cada una de las veredas.
Sector	Numérico	No	No	Número que identifica el sector como punto de referencia dentro de la vereda.
Tipo1	Texto	No	No	Nombre de la entidad en primera instancia que atiende el caso.
Tipo2	Texto	Si	No	Nombre de la entidad que atiende el caso en segunda instancia si se requiere.
Tipo3	Texto	Si	No	Nombre de la entidad que atiende el caso en tercera instancia si se requiere.
Otro	Texto	Si	No	Nombre de otra entidad que atiende el caso si se requiere.
Tipoe	Texto	No	No	Tipificación de las emergencias reportadas por el usuario.
Descrip	Texto	No	No	Descripción o ampliación de los hechos de la emergencia reportada.
Nombreq	Texto	No	No	Nombre del funcionario que recibe la llamada.
Entidad	Texto	No	No	Nombre de la entidad a la que pertenece el funcionario quien recibe la llamada.
Descrip2	Texto	Si	No	Notas de ampliación del caso.

Segui	Texto	No	No	Referencia si el caso requiere o no seguimiento.
Seguinombreuser	Texto	No	No	Nombre del funcionario que realiza el seguimiento al caso.
Seguidescrip	Texto	No	No	Detalles de lo transcurrido en el seguimiento.
Seguifecha	Numérico	Si	No	Fecha de los detalles que transcurrieron dentro del seguimiento del caso.
Cerrado	Texto	No	No	Estado final del caso si se concluyó o no.

Llave primaria (Si aplica)	
Llave(s) Foránea(s)	
Otras Restricciones	

IDENTIFICACIÓN DE DATOS Fuentes de Extracción	Código: 2501
	Versión: V2
	Página: 1
	Vigente: Si

1. IDENTIFICACIÓN DE LOS DATOS

1.1 Nombre de la Fuente	EmergenciasFinal.xlsx					
1.2 Descripción detallada de la fuente datos (Detalle a que hacen referencia los datos)	Datos proporcionados por la aplicación de recepción de llamadas de la línea de emergencia 123 del Municipio de Chía.					
1.3 Tipo de Origen de los Datos	WEB		APLICACION	X	BD	OTRO
1.4 URL o Referencia de origen						
1.5 Cantidad de Variables						
1.6 Cantidad de Registros	5					
1.5 Peso físico – Si aplica	683 Kb					

2. DICCIONARIO DE DATOS

Nombre Campo	Tipo de Dato	Nulo (S/N)	Transformación (S/N)	Descripción
ID_EMERGENCIA	Numérico	No	No	Consecutivo generado automáticamente por el aplicativo correspondiente a la identificación de la llamada.
ID_USUARIO	Numérico	No	No	Consecutivo generado automáticamente por el aplicativo correspondiente a la

				identificación del usuario que realizo la llamada.
LATITUD	Numérico	No	No	Coordenadas de georreferenciación de la ubicación del evento reportado en la llamada.
LONGITUD	Numérico	No	No	Coordenadas de georreferenciación de la ubicación del evento reportado en la llamada.
FECHA	Numérico	No	No	Fecha y hora generada por el aplicativo correspondiente al momento de la recepción de la llamada.

Llave primaria (Si aplica)	
Llave(s) Foránea(s)	
Otras Restricciones	

Informe Policial de Accidentes de Tránsito (IPAT) - Secretaría de Movilidad.

Origen de los datos: Recopilación manual por parte de los agentes de tránsito durante la atención de accidentes.

Formato: Formulario físico establecido por la Resolución 11268 de 2012 del Ministerio de Transporte.

Contenido: Información detallada sobre el accidente, incluyendo datos de los involucrados, vehículos, lugar, hora, condiciones de tránsito, entre otros.

Proceso de captura: Los agentes diligencian el formulario en el lugar del accidente y posteriormente, un funcionario transcribe los datos a una base de datos local.

Frecuencia de actualización: Cada vez que se atiende un accidente de tránsito.

Limitaciones: Posibilidad de errores humanos durante la transcripción, omisión de datos o registros incompletos

IDENTIFICACIÓN DE DATOS Fuentes de Extracción	Código: 2501
	Versión: V2
	Página: 1
	Vigente: Si

1. IDENTIFICACIÓN DE LOS DATOS

1.1 Nombre de la Fuente	MATRIZ IPAT						
1.2 Descripción detallada de la fuente datos (Detalle a que hacen referencia los datos)							
1.3 Tipo de Origen de los Datos	WEB		APLICACION		BD		OTRO X
1.4 URL o Referencia de origen							
1.5 Cantidad de Variables							
1.6 Cantidad de Registros	131						
1.5 Peso físico – Si aplica							

2. DICCIONARIO DE DATOS

Nombre Campo	Tipo de Dato	Nulo (S/N)	Transformación (S/N)	Descripción
Nº	Numérico	No	No	Campo de tipo numérico consecutivo con el fin de identificar en orden el número de casos relacionados.
Nº IPAT	Alfanumérico	No	No	Campo de tipo alfanumérico de manejo interno de la entidad, que representa "Informe Policial de Accidentes de Tránsito".

Gravedad	Texto	No	No	Campo que indica la característica involucrada en el tipo de accidente ocurrido.
Nombre vía principal.	Alfanumérico	No	No	Campo para la dirección o descripción del lugar donde ocurrió el accidente.
Latitud	Numérico	Si	No	Punto de georreferencia determinado en grados minutos y segundos, dato no indispensable.
Longitud	Numérico	Si	No	Punto de georreferencia determinado en grados minutos y segundos, dato no indispensable.
Localidad	Texto	No	No	Dirección o punto de referencia del sitio correspondiente al accidente.
Día	Numérico	No	No	Número correspondiente al día del accidente.
Mes	Numérico	No	No	Número correspondiente al mes del accidente.
Año	Numérico	No	No	Número correspondiente al mes del accidente.
Hora	Numérico	No	No	Número correspondiente a la hora de los hechos del accidente.
Minutos	Numérico	No	No	Número correspondiente a los minutos de los hechos del accidente.
Día de la semana	Texto	No	No	Texto correspondiente al día de la semana cuando ocurrieron los hechos.
Clase de accidente	Texto	No	No	Relación del tipo de accidente ocurrido.
Choque con	Texto	Si	No	Objeto que hace referencia a con que se estrelló un vehículo.
Objeto fijo	Texto	Si	No	Descripción del objeto con el cual hubo el accidente.
Área	Texto	No	No	Descripción geográfica de la zona donde ocurrió el accidente.
Sector	Texto	No	No	Tipo de actividad económica desarrollada en el sector donde ocurrieron los hechos.

Zona	Texto	No	No	Tipo de zona que predomina en el sector donde ocurrieron los hechos.
Diseño	Texto	No	No	Hace referencia a las especificaciones de la vía donde ocurrieron los hechos.
Condición climática	Texto	No	No	Estado del clima en el momento de los Hechos.
Geometría	Texto	No	No	Estructura de la vía donde ocurrió el accidente.
Utilización	Texto	No	No	Indicación del Sentido en el que transitan los vehículos en el lugar donde ocurrieron los hechos.
Calzadas	Texto	No	No	Numero de zonas donde transitan los vehículos en el lugar del accidente.
Carriles	Texto	No	No	Número de carriles de la calzada donde se generaron los hechos.
Superficie	Texto	No	No	Material de construcción que predomina en la vía donde ocurrieron los hechos.
Estado	Texto	No	No	Detalle del estado físico de la vía donde ocurrieron los hechos.
Condiciones	Texto	No	No	factores climáticos que afectaron la vía en el momento del accidente.
Iluminación artificial	Texto	Si	No	Determina la condición del alumbrado público en el sector del accidente.
Agente de transito	Texto	No	No	Determina la presencia de un agente de tránsito en el sitio del accidente.
Semáforo	Texto	No	No	Indicar si hay semaforización en el lugar del accidente y su estado de funcionamiento.
Señal vertical	Texto	Si	No	Determinar si hay o no señalización en el sitio y a cual corresponde.
Señal horizontal	Texto	Si	No	Determinar si hay o no señalización en el sitio y a cual corresponde.

Reductor de velocidad	Texto	Si	No	Determinar si hay o no reductor en el sitio y a cuál corresponde.
Delineación de piso	Texto	Si	No	Determinar si hay o no delineación en el sitio y a cuál corresponde.
Visibilidad	Texto	No	No	Indica el estado de visibilidad al momento del accidente.
Código	Numérico	No	No	Código interno que identifica y relaciona la hipótesis
Hipótesis vehículo	Texto	No	No	Indicación del conductor según su criterio de los hechos se relaciona a la tabla de hipótesis
Aclaración hipótesis	Texto	No	No	Corroboración de la hipótesis relacionada
Entidad	Texto	No	No	Nombre de la entidad y código del agente que atendió el caso
Nombre	Texto	No	No	Nombre del funcionario encargado de la atención del caso
Inmovilización	Texto	No	No	Determina si hubo o no inmovilización de alguno de los vehículos involucrados.
Observaciones	Texto	Si	No	Relaciona alguna observación por parte de la entidad que atiende el caso
Funcionario que digita	Texto	No	No	Nombre del funcionario encargado de digitalizar el proceso.
Conductor	Texto	No	No	Nombre completo del conductor de cada vehículo
Tipo de documento	Texto	No	No	Referenciar tipo de documento que presenta el conductor
Numero	Numérico	No	No	Número de identificación de los conductores
Nacionalidad	Texto	No	No	Lugar de nacionalidad de los conductores
Gravedad	Texto	No	No	Indicar la consecuencia del accidente
Lugar de domicilio	Alfanumérico	No	No	Dirección de residencia de los conductores involucrados

Genero	Texto	No	No	Tipo de genero ya sea masculino o femenino que idéntica al conductor.
Fecha de nacimiento	Numérico	No	No	Día ,mes y año de nacimiento del conductor
Edad	Numérico	No	No	Edad determinada por el conductor
Embriaguez	Texto	No	No	Indicar si el conductor se encuentra en estado de embriaguez al momento del accidente
Placa	Alfanumérico	No	No	Dato de la placa del vehículo involucrado
Clase de vehículo	Texto	No	No	Determinar qué clase de vehículo es el involucrado, ya sea camión, automóvil, motocicleta u otro
Soat	Texto	No	No	Determinar si el vehículo involucrado al momento del accidente cuenta con soat vigente
Técnico mecánica	Texto	No	No	Determinar si el vehículo involucrado al momento del accidente cuenta con revisión técnico mecánica vigente
Clase de servicio	Texto	No	No	Determinar si el vehículo involucrado es de clase particular, publico u otro
Modalidad de transporte	Texto	No	No	Indicar si el vehículo corresponde a modalidad de carga, colectivo, individual o mixto.
Radio de acción	Texto	No	No	Determina si el vehículo procede en un radio de acción municipal o nacional
Vehículo	Numérico	Si	No	Relación del vehículo de donde resultado el herido
Nombre	Texto	No	No	Nombre completo del herido
Tipo documento	Texto	No	No	Tipo de documento de identificación del herido
Numero	Numérico	No	No	Número de identificación del herido

Fecha de nacimiento	Numérico	No	No	Dato de la fecha de nacimiento del herido que incluye día, mes y año
Edad	Numérico	No	No	Dato de los años del herido
Condición	Texto	No	No	Rol que cumple el herido dentro del accidente, ya se acompañante, peatón u otro
Gravedad	Texto	No	No	Estado del herido al momento del accidente
Lugar de domicilio	Alfanumérico	No	No	Dirección de residencia del herido

Llave primaria (Si aplica)	
Llave(s) Foránea(s)	
Otras Restricciones	

IDENTIFICACIÓN DE DATOS Fuentes de Extracción	Código: 2501
	Versión: V2
	Página: 1
	Vigente: Si

1. IDENTIFICACIÓN DE LOS DATOS

1.1 Nombre de la Fuente	MATRIZ IPAT							
1.2 Descripción detallada de la fuente datos (Detalle a que hacen referencia los datos)	Datos relacionados de los conductores involucrados en el accidente.							
1.3 Tipo de Origen de los Datos	WEB		APLICACION		BD		OTRO	X
1.4 URL o Referencia de origen								
1.5 Cantidad de Variables								
1.6 Cantidad de Registros	19							
1.5 Peso físico – Si aplica	23 KB							

2. DICCIONARIO DE DATOS

Nombre Campo	Tipo de Dato	Nulo (S/N)	Transformación (S/N)	Descripción
Nº	Numérico	No	No	Campo de tipo numérico consecutivo con el fin de identificar en orden el número de casos relacionados.
Nº IPAT	Alfanumérico	No	No	Campo de tipo alfanumérico de manejo interno de la entidad, que representa "Informe Policial de Accidentes de Tránsito".

Conductor	Texto	No	No	Nombre completo del conductor de cada vehículo
Tipo de documento	Texto	No	No	Referenciar tipo de documento que presenta el conductor
Numero	Numérico	No	No	Número de identificación de los conductores
Nacionalidad	Texto	No	No	Lugar de nacionalidad de los conductores
Gravedad	Texto	No	No	Indicar la consecuencia del accidente
Lugar de domicilio	Alfanumérico	No	No	Dirección de residencia de los conductores involucrados
Genero	Texto	No	No	Tipo de genero ya sea masculino o femenino que idéntica al conductor.
Fecha de nacimiento	Numérico	No	No	Día ,mes y año de nacimiento del conductor
Edad	Numérico	No	No	Edad determinada por el conductor
Embriaguez	Texto	No	No	Indicar si el conductor se encuentra en estado de embriaguez al momento del accidente
Placa	Alfanumérico	No	No	Dato de la placa del vehículo involucrado
Clase de vehículo	Texto	No	No	Determinar qué clase de vehículo es el involucrado, ya sea camión, automóvil, motocicleta u otro
Soat	Texto	No	No	Determinar si el vehículo involucrado al momento del accidente cuenta con soat vigente
Técnico mecánica	Texto	No	No	Determinar si el vehículo involucrado al momento del accidente cuenta con revisión técnico mecánica vigente
Clase de servicio	Texto	No	No	Determinar si el vehículo involucrado es de clase particular, publico u otro
Modalidad de transporte	Texto	No	No	Indicar si el vehículo corresponde a modalidad de carga, colectivo, individual o mixto.

Radio de acción	Texto	No	No	Determina si el vehículo procede en un radio de acción municipal o nacional
Vehículo	Numérico	Si	No	Relación del vehículo de donde resulto el herido
Nombre	Texto	No	No	Nombre completo del herido
Tipo documento	Texto	No	No	Tipo de documento de identificación del herido
Numero	Numérico	No	No	Número de identificación del herido
Fecha de nacimiento	Numérico	No	No	Dato de la fecha de nacimiento del herido que incluye día, mes y año
Edad	Numérico	No	No	Dato de los años del herido
Condición	Texto	No	No	Rol que cumple el herido dentro del accidente, ya se acompañante, peatón u otro
Gravedad	Texto	No	No	Estado del herido al momento del accidente
Lugar de domicilio	Alfanumérico	No	No	Dirección de residencia del herido

Llave primaria (Si aplica)	
Llave(s) Foránea(s)	
Otras Restricciones	

IDENTIFICACIÓN DE DATOS Fuentes de Extracción	Código: 2501
	Versión: V2
	Página: 1
	Vigente: Si

1. IDENTIFICACIÓN DE LOS DATOS

1.1 Nombre de la Fuente	MATRIZ IPAT							
1.2 Descripción detallada de la fuente datos (Detalle a que hacen referencia los datos)	Datos relacionados de los conductores involucrados en el accidente.							
1.3 Tipo de Origen de los Datos	WEB		APLICACION		BD		OTRO	X
1.4 URL o Referencia de origen								
1.5 Cantidad de Variables								
1.6 Cantidad de Registros	11							
1.5 Peso físico – Si aplica	24 KB							

2. DICCIONARIO DE DATOS

Nombre Campo	Tipo de Dato	Nulo (S/N)	Transformación (S/N)	Descripción
Nº	Numérico	No	No	Campo de tipo numérico consecutivo con el fin de identificar en orden el número de casos relacionados.
Nº IPAT	Alfanumérico	No	No	Campo de tipo alfanumérico de manejo interno de la entidad, que representa "Informe Policial de Accidentes de Tránsito".

Vehículo	Numérico	Si	No	Relación del vehículo de donde resulto el herido
Nombre	Texto	No	No	Nombre completo del herido
Tipo documento	Texto	No	No	Tipo de documento de identificación del herido
Numero	Numérico	No	No	Número de identificación del herido
Fecha de nacimiento	Numérico	No	No	Dato de la fecha de nacimiento del herido que incluye día, mes y año
Edad	Numérico	No	No	Dato de los años del herido
Condición	Texto	No	No	Rol que cumple el herido dentro del accidente, ya se acompañante, peatón u otro
Gravedad	Texto	No	No	Estado del herido al momento del accidente
Lugar de domicilio	Alfanumérico	No	No	Dirección de residencia del herido

Llave primaria (Si aplica)	
Llave(s) Foránea(s)	
Otras Restricciones	

EJECUCIÓN METODOLÓGICA

1. PREPARACIÓN DE LOS DATOS

1.1 Importancia estratégica de la preparación de datos

Dentro de la metodología CRISP-DM, la fase de Preparación de los Datos representa el núcleo técnico del proyecto de analítica. La calidad de los análisis, la confiabilidad de los resultados y la sostenibilidad del observatorio dependen directamente del rigor aplicado en esta etapa.

En el contexto institucional del Municipio de Chía, la preparación de datos adquiere una relevancia superior debido a la heterogeneidad de las fuentes, la coexistencia de procesos manuales y automatizados, y la necesidad de garantizar trazabilidad, integridad y reutilización de la información a lo largo del tiempo.

1.2 Arquitectura general del proceso ETL

El proceso ETL (Extract, Transform, Load) se diseña como una arquitectura lógica que permite el flujo controlado de los datos desde las fuentes operacionales hasta los repositorios analíticos.

Esta arquitectura contempla:

- Capas de extracción por tipo de fuente.
- Capas de transformación con reglas explícitas.
- Capas de carga orientadas al análisis.
- Mecanismos de control y auditoría.

El diseño modular del ETL facilita su evolución futura y su integración con nuevas fuentes de información.

1.3 Extracción de datos

La extracción de datos se realiza desde múltiples orígenes, entre los que se incluyen sistemas automáticos como el aplicativo de la línea de emergencias 123 y fuentes manuales como los formularios IPAT.

Cada fuente es caracterizada previamente mediante:

- Identificación del origen.
- Tipo de captura.
- Periodicidad.
- Volumen estimado.
- Nivel de confiabilidad.

Este proceso permite establecer estrategias diferenciadas de tratamiento según el origen de la información.

1.4 Transformación de datos

La transformación de los datos es una de las actividades más críticas del proceso ETL. En esta etapa se ejecutan múltiples operaciones destinadas a mejorar la calidad y coherencia de la información.

Las transformaciones incluyen:

- Conversión de tipos de datos.
- Homologación de valores categóricos.
- Normalización de formatos.

- Generación de variables analíticas.
- Aplicación de reglas de negocio.

Cada transformación es documentada para asegurar trazabilidad técnica.

1.5 Reglas de limpieza de datos

Las reglas de limpieza se definen con el objetivo de mitigar errores inherentes a la captura manual y a la heterogeneidad de los sistemas.

Entre las principales reglas se incluyen:

- Identificación y eliminación de duplicados.
- Tratamiento de valores nulos.
- Validación de rangos aceptables.
- Corrección de inconsistencias.
- Control de valores atípicos.

La aplicación sistemática de estas reglas incrementa significativamente la confiabilidad del análisis.

1.6 Normalización y estandarización

La normalización de datos permite integrar información proveniente de múltiples fuentes bajo una estructura homogénea. Este proceso es esencial para garantizar comparabilidad y consistencia.

Se normalizan:

- Fechas y horas.
- Códigos geográficos.
- Tipologías de eventos.
- Clasificaciones institucionales.

La estandarización reduce la ambigüedad semántica y facilita el análisis transversal.

1.7 Variables derivadas y enriquecimiento

El enriquecimiento de los datos mediante variables derivadas permite ampliar el valor analítico de la información.

Se generan variables como:

- Franja horaria.
- Día de la semana.

- Periodo del año.
- Clasificación geográfica.
- Indicadores de severidad.

Estas variables permiten análisis más profundos y orientados a la toma de decisiones.

2. MODELADO

2.1 Enfoque general del modelado

El modelado analítico se orienta a comprender el comportamiento de los datos y a identificar patrones relevantes. Dado el nivel de madurez de la información, se priorizan modelos descriptivos y diagnósticos.

Este enfoque garantiza resultados interpretables, útiles y alineados con las capacidades institucionales.

2.2 Análisis Exploratorio de Datos (EDA)

El EDA constituye el primer acercamiento analítico a los datos preparados. Su objetivo es comprender la distribución, variabilidad y relaciones entre las variables.

Se utilizan:

- Estadísticas descriptivas.
- Análisis de frecuencia.
- Visualizaciones exploratorias.
- Identificación de anomalías.

El EDA orienta decisiones posteriores de análisis y segmentación.

2.3 Análisis espacial

El análisis espacial permite identificar concentraciones geográficas de eventos y zonas críticas.

Se apoya en información georreferenciada y en divisiones administrativas del municipio.

Este análisis facilita la priorización territorial de intervenciones y recursos.

2.4 Análisis temporal

El análisis temporal evalúa la evolución de los eventos en el tiempo. Permite identificar tendencias, estacionalidades y patrones horarios.

Este enfoque es clave para la planificación operativa y preventiva.

2.5 Segmentación analítica

La segmentación clasifica los eventos en grupos homogéneos según criterios definidos.

Facilita la focalización de estrategias y la comunicación de resultados.

2.6 Límites y restricciones del modelado

El alcance del modelado se limita deliberadamente para evitar sobreinterpretaciones.

No se aplican modelos predictivos avanzados hasta fortalecer la calidad y continuidad de los datos.

3. EVALUACIÓN

3.1 Evaluación técnica

La evaluación técnica verifica la consistencia y reproducibilidad de los resultados analíticos.

Incluye revisión de procesos ETL, estabilidad de resultados y trazabilidad.

3.2 Evaluación funcional

La evaluación funcional valida la utilidad de los resultados para las dependencias usuarias.

Este proceso asegura alineación con la realidad operativa.

3.3 Criterios de aceptación

Los criterios de aceptación establecen condiciones claras para la validación del proceso analítico, incluyendo utilidad, coherencia y alineación institucional.

4. DESPLIEGUE

4.1 Operación del observatorio

El despliegue operativo materializa los resultados analíticos en productos utilizables, como informes y tableros de control.

4.2 Actualización y mantenimiento

Se definen lineamientos para la actualización periódica de los datos y la sostenibilidad técnica del proceso.

4.3 Sostenibilidad institucional

La sostenibilidad depende de documentación, capacitación, gobierno de datos y compromiso institucional.

5. ALINEACIÓN CON ISO/IEC 27001

5.1 Seguridad de la información como eje transversal

La información analizada constituye un activo crítico del municipio y debe ser protegida bajo principios de confidencialidad, integridad y disponibilidad.

5.2 Controles de seguridad aplicables

Se identifican controles alineados con ISO/IEC 27001 que fortalecen la protección de la información analítica.

5.3 Gestión de riesgos de información

Se identifican riesgos asociados a la manipulación, almacenamiento y uso de los datos, definiendo controles técnicos y procedimentales.

MATRICES POR FASES

1.2 Matriz ETL por Fuente de Datos

Fuente	Campo Origen	Regla Limpieza	Transformación	Campo Destino	Observatorio
Línea 123	Fecha	Validar formato	Normalizar datetime	fecha_evento	Seguridad
Línea 123	Zona	No nulo	Homologar catálogo	zona_id	Seguridad
IPAT	Hora	Rango válido	Franja horaria	franja_horaria	Movilidad
IPAT	Clase accidente	Texto estándar	Codificación	tipo_accidente	Movilidad

La matriz ETL documenta de forma explícita las reglas aplicadas a cada campo, garantizando trazabilidad técnica y facilitando procesos de auditoría y control institucional.

1.3 Reglas de Calidad de Datos (DQ Rules)

Dimensión	Descripción	Aplicación 123	Aplicación IPAT
Compleitud	Campos obligatorios diligenciados	Alta	Media
Consistencia	Valores coherentes	Alta	Media
Validez	Rangos y formatos correctos	Alta	Media
Unicidad	Registros no duplicados	Alta	Baja
Oportunidad	Actualización oportuna	Tiempo real	Diferido

2. MODELADO

El modelado se orienta a análisis descriptivo y diagnóstico, alineado con la madurez actual de los datos municipales. Se prioriza la interpretabilidad y el valor institucional sobre la complejidad algorítmica.

2.1 Catálogo de Análisis Aplicados y Potenciales

Tipo de Análisis	Descripción	Observatorio	Estado
EDA	Distribuciones y frecuencias	Ambos	Implementado
Análisis espacial	Zonas críticas	Ambos	Implementado
Análisis temporal	Tendencias	Ambos	Implementado
Clustering	Hotspots	Movilidad	Potencial
Series de tiempo	Predicción básica	Seguridad	Potencial

2.2 Hipótesis Analíticas

Hipótesis	Variable Dependiente	Variables Independientes	Uso Institucional
H1	Severidad accidente	Hora, zona, clima	Prevención vial
H2	Frecuencia llamadas	Zona, franja horaria	Planeación seguridad

6. EVALUACIÓN

La evaluación valida que los resultados analíticos sean técnicamente correctos y funcionalmente útiles para las dependencias del Municipio de Chía.

3.1 Matriz de Criterios de Aceptación

Criterio	Descripción	Responsable	Evidencia
Coherencia	Consistencia de resultados	TIC	Informe técnico
Utilidad	Apoya decisiones	Dependencia	Acta validación
Trazabilidad	Origen del dato claro	TIC	Diccionario

4. DESPLIEGUE

El despliegue operacionaliza los resultados mediante productos analíticos reutilizables y sostenibles en el tiempo.

4.1 Modelo Operativo del Observatorio

Rol	Responsabilidad	Periodicidad
Analista	Actualización dashboards	Mensual
TIC	Soporte técnico	Permanente
Dependencia	Uso del indicador	Continuo

5. ROADMAP DE MADUREZ ANALÍTICA

El roadmap define la evolución progresiva del uso de la analítica de datos en los observatorios.

Nivel	Tipo Analítica	Seguridad	Movilidad
1	Descriptiva	Indicadores básicos	Indicadores básicos
2	Diagnóstica	Causas y patrones	Zonas críticas
3	Predictiva	Tendencias	Riesgo vial
4	Prescriptiva	Recomendaciones	Optimización

6. ALINEACIÓN ISO/IEC 27001

La información gestionada por los observatorios se reconoce como un activo crítico y se alinea con los principios de la norma ISO/IEC 27001.

6.1 Inventario de Activos de Información

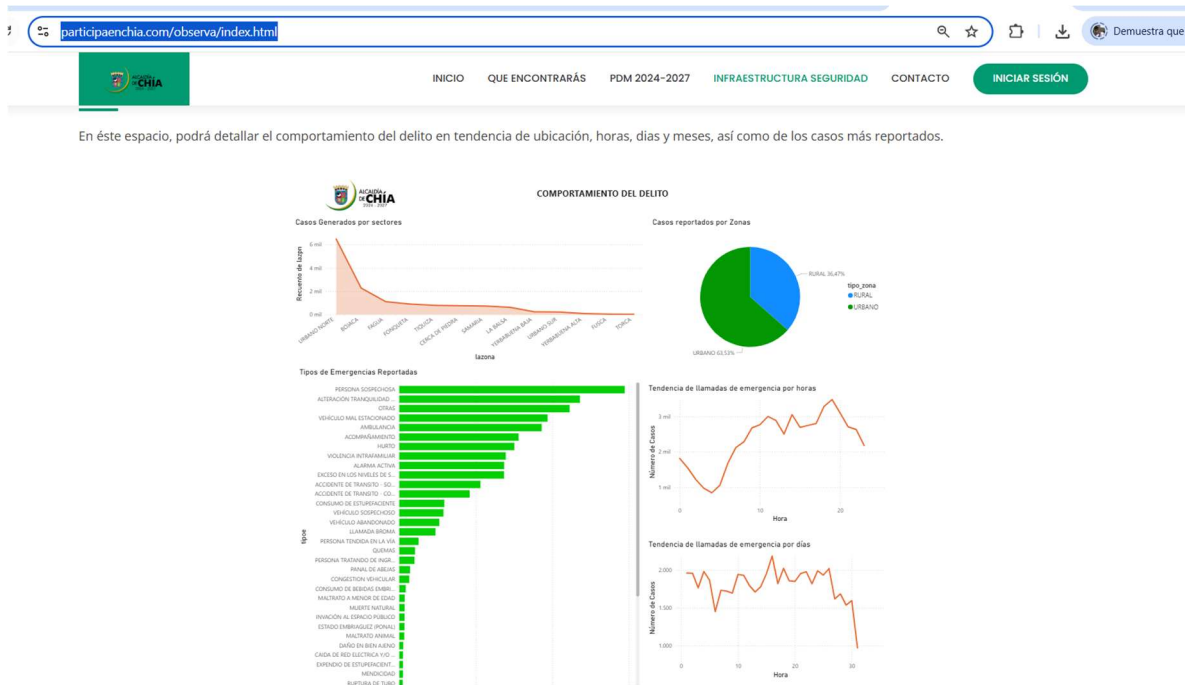
Activo	Observatorio	Clasificación	Responsable
BD 123	Seguridad	Confidencial	Dirección Seguridad
Matriz IPAT	Movilidad	Confidencial	Sec. Movilidad

6.2 Matriz de Riesgos

Activo	Amenaza	Impacto	Riesgo	Control
BD 123	Acceso no autorizado	Alto	Alto	Control accesos
IPAT	Error digitación	Medio	Medio	Validación

EVIDENCIA DE PUBLICACIÓN

<https://participaenchia.com/observa/index.html>



2. Firmas

2.1 Responsable Proceso:

Ingeniero de Desarrollo

2.2 VoBo.

Ingeniero de Desarrollo