



TDSP COLISIONES VEHÍCULOS PERSONAS

Minería de Datos

Integrantes

Ángel Fabricio Zacarías Guzmán
Lilián Hernández Contreras
Leonardo Manuel Medina Rodríguez

Problemática:

Dentro de la ciudad de Nueva York, se han presentado una innumerable cantidad de casos por choques automovilísticos los cuales exceden de daños físicos, materiales, entre otros que superan los \$1,000 dólares. Todos los datos que se muestran han sido registrados por los Agentes de Policía de Nueva York (NYPD por sus siglas en inglés).

Esto ha llegado a causar una gran intriga por parte de las autoridades, ya que los casos se encuentran cada vez más en aumento, y así mismo se quiere saber cuáles son las horas clave para realizar un análisis de cómo evitar muchas de estas grandes tragedias.

Propuesta Solución:

Por lo que se propone, el desarrollo de una limpieza de datos, análisis y entendimiento de ellos. Para a continuación crear un modelo de "Machine Learning", haciendo su correspondiente validación de los Modelos de Prueba. Una vez finalizado se espera mostrar una representación gráfica utilizando herramientas como Tablas y Gráficas.

Objetivos:

General:

- Desarrollar un sistema el cual permita a las autoridades de Nueva York visualizar de una mejor manera la problemática que se está teniendo y buscar acciones de tránsito para evitar muchos de estos conflictos a futuro.

Específicos:

- Apoyar de una manera directa a la policía de Nueva York para la toma de decisiones.
- Bajar la cantidad de choques automovilísticos.
- Ayudar a los dueños de automóviles a crear consciencia sobre sus hábitos de manejo.
- Ayudar a los transeúntes que se encuentra involucrados en los accidentes a contar con más protecciones al momento de viajar.

Alcance:

El alcance inicial se propone inicialmente para la Ciudad de Nueva York (Estados Unidos), siendo así de las más grandes metrópolis del mundo, ayudará a después pasar la información a cualquier lugar de nuestro planeta tierra para mejorar las condiciones de tránsito de las ciudades.

Limitaciones:

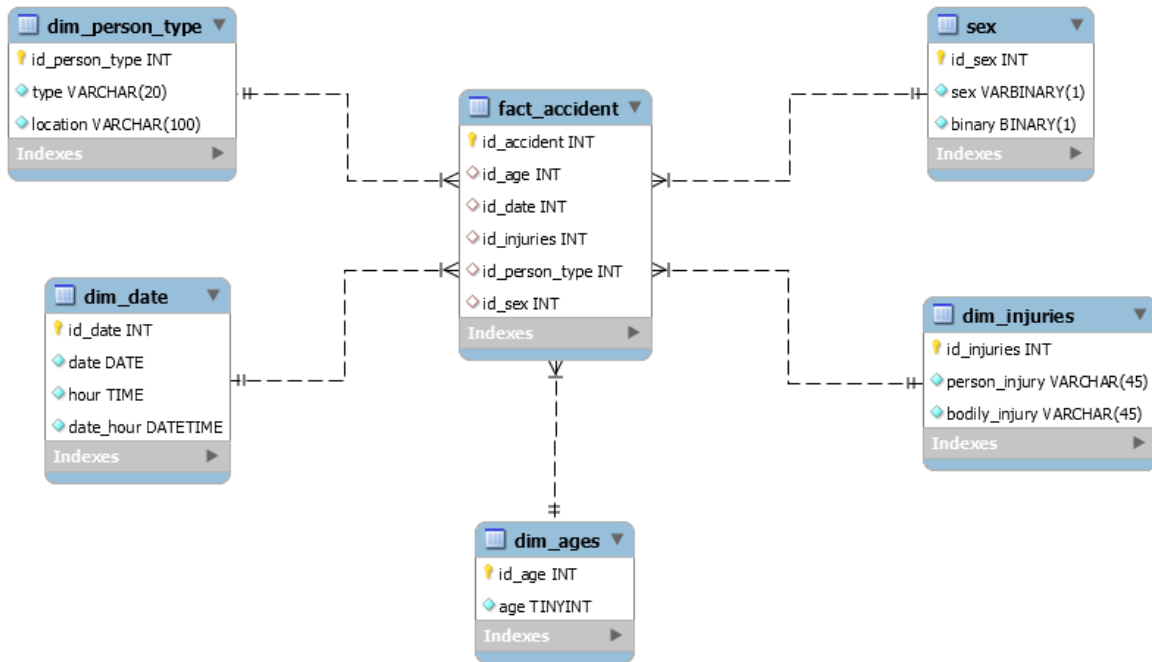
Siendo una de las grandes limitaciones es la falta de información correcta que registren las unidades de policía, ya que se puede tener una diferencia de datos al momento de ellos registrarlo.

Así mismo estos datos se van especificando cada cierto tiempo o se hacen ya sea de manera anual o mensual. Lo cual no se tiene una fuente continua de datos para llevar un análisis al momento del registro de éstos.

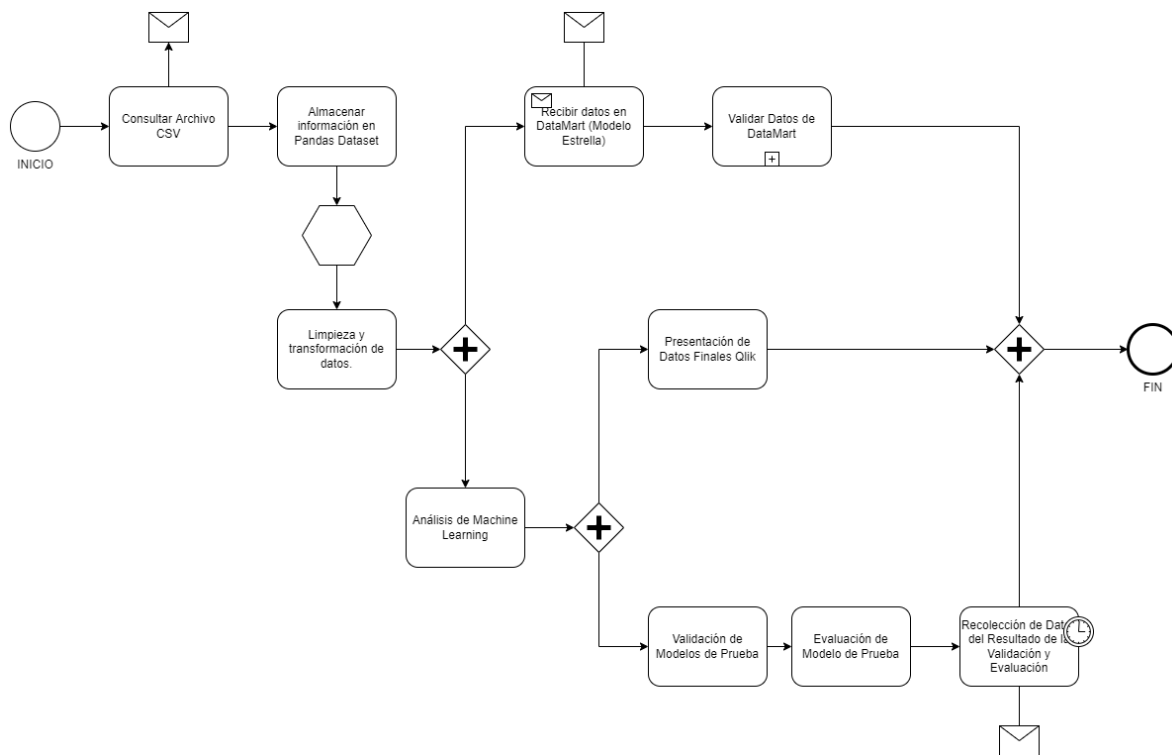
Diagramas

Entidad Relación DataMart

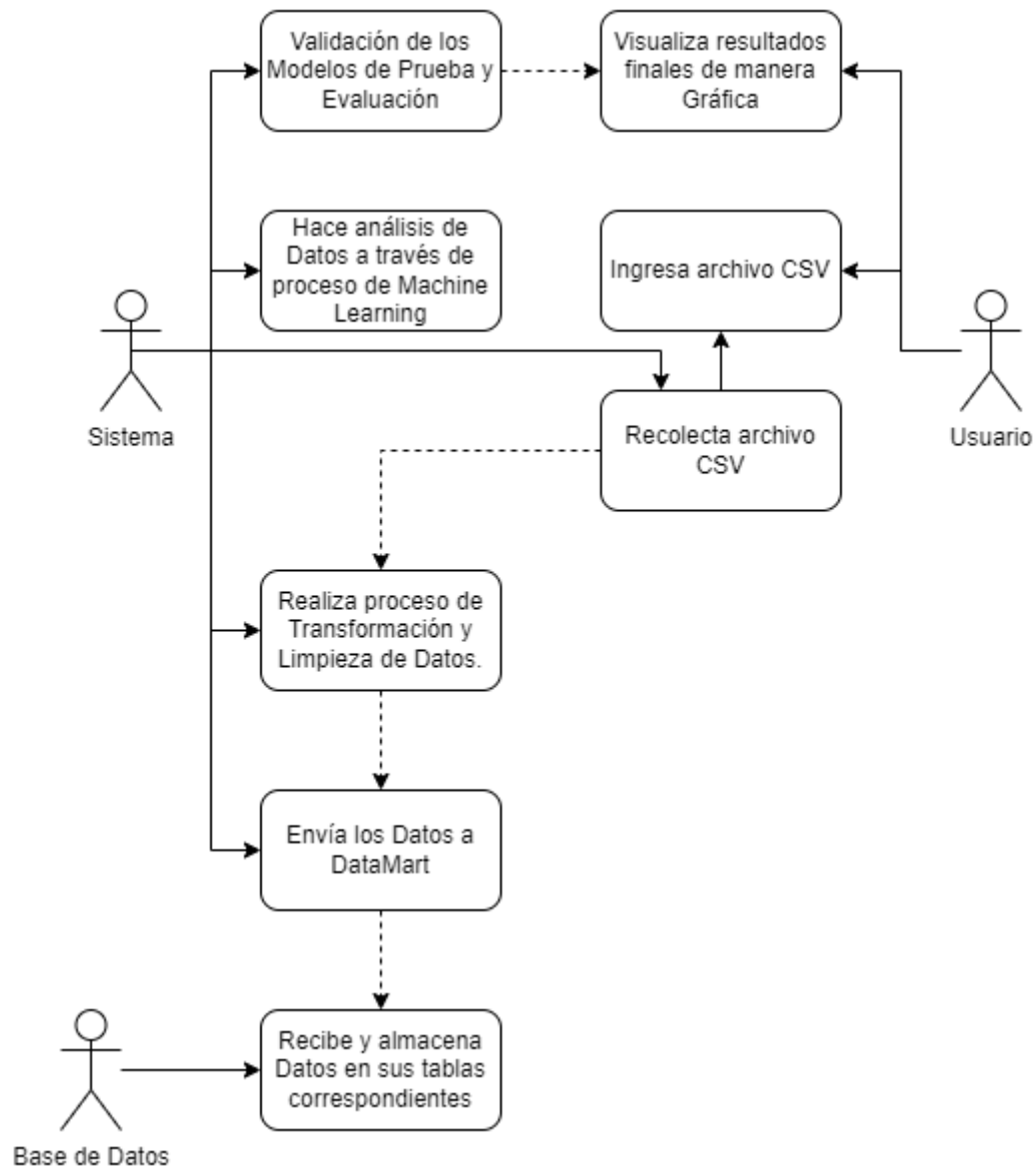
- Se realiza el Modelo Estrella en el diagrama entidad relación para el guardado de la información, siendo la tabla de “Hechos” la de tabla de accidentes.



BPMN



Casos de Uso



Requerimientos de Sistema

Requerimientos Funcionales

Requerimiento	descripción	Tipo
RF01	Recolección de datos: El software recolectara datos previos, que fueron añadidos por el usuario.	Sistema
RF02	Limpieza de datos: El software limpiara los datos en caso de que exista algún dato vacío o nulo.	Sistema
RF03	transformación de datos: El software se encargará de transformar los datos almacenados, en caso de que se requiera esta acción.	Sistema
RF04	normalización de datos: El software se encargará de normalizar los datos ingresados, para tener una mejor organización.	Sistema
RF05	El usuario asignara los valores y los procedimientos que el software utilizara para su correcto funcionamiento.	Usuario
RF06	El usuario puede asignar el documento que sea de su elección.	Usuario
RF07	El software debe contar con gráficos que muestren los resultados obtenidos.	Sistema
RF08	El software debe mostrar la información del documento en su interfaz.	Sistema
RF09	El software modelara un algoritmo para su uso.	Sistema
RF10	El software validará un modelo de prueba.	Sistema
RF11	El software validará un modelo de evaluación.	Sistema
RF12	El software debe obtener toda su información principal de un documento que asigne el usuario.	Sistema
RF13	El usuario se encargará de buscar el problema que solucionará el software	Usuario
RF14	DataMart: El software validará los datos que se encuentran almacenados en un DataMart.	Sistema
RF15	El software mostrara los datos previos a ser modificados, así como los datos que sean modificados.	Sistema

Requerimientos No Funcionales

Requerimiento	descripción	Tipo
RNF01	El software necesitara un documento que almacene más de 1500 campos de información.	Sistema
RNF02	El software debe ser fácil de usar.	Sistema
RNF03	El software contará con un manual de uso.	Sistema
RNF04	El software debe ser seguro. (Confidencialidad en datos).	Sistema
RNF05	Los datos del software deben estar almacenados sobre un gestor de base de datos.	Sistema
RNF06	La base de datos debe contener con un gran espacio de almacenamiento.	Sistema
RNF07	El software debe funcionar el 99% del tiempo.	Sistema
RNF08	El software debe tener un tiempo de respuesta razonable.	Sistema
RNF09	La documentación técnica se entregará en formato digital.	Sistema
RNF10	Se debe utilizar un sistema operativo de Microsoft.	Sistema
RNF11	Debe ser desarrollado con el ide Jupyter (Anaconda)	Sistema
RNF12	Se debe contar con una licencia de Microsoft.	Sistema
RNF13	El documento debe ser referente a una empresa real.	Sistema
RNF14	El software debe tener interfaces llamativas.	Sistema
RNF15	El manual del software se entregará en formato digital.	Sistema