

1.

Se podría preguntar lo siguiente:

¿Qué información específica tiene sobre los siniestros?

Por ejemplo, la hora, podría ser la ubicación

¿Cómo han sucedido los accidentes, de qué manera?

¿Cómo se almacenan los tipos de autos en la póliza?

¿Qué información contienen de los clientes asegurados?

Historial de conducción de los asegurados para que se descarte que los usuarios o algunos conducen mal o en estado de ebriedad

Variables categóricas:

Todas las relacionadas con el vehículo, las que menciona el texto como **tipo de auto, marca, modelo, año de compra, tipo de uso.**

Ubicación del Siniestro

Perfil de cliente

Considero que con esta información ayudamos al análisis de los patrones de siniestros y posibles causas de su aumento.

2.

*****Pregunta respondida en la hoja.**

3.

Plan de Trabajo

Definición del Problema: Entender el problema de siniestros y ha existido un incremento en el último año.

Recolección de Datos: Obtener datos sobre pólizas y siniestros con la información que se recabó en la pregunta 1.

Preparación de Datos: Limpiar, organizar los datos y utilizar los más apropiados o los que servirán para el análisis de datos.

Análisis Exploratorio: Buscar patrones y tendencias o eventualidades parecidas y cuáles fueron sus posibles causas para desde ahí poder partir.

Modelado de Datos: Crear y probar modelos predictivos.

Implementación: Aplicar soluciones y recomendaciones.

Monitoreo: Seguir los resultados y ajustar según sea necesario para corregir o probar nuevas soluciones.

4.

Mayor confianza en el cliente

Para que le cliente tenga mayor confianza,

Transparencia: para que su información sepa el cómo es utilizada y para que fines será utilizada y que él pueda documentar esa información y corroborar que es correcta esa utilización de su información.

Experiencia: resaltar la experiencia y conocimientos y manejo de la información por parte del equipo de Minería de Datos, para que el vea que somos capaces de solucionar su problema.

Avances: Dar avances o reportes de trabajo realizado y si existe un resultado con éxito, dárselo a la brevedad y él pueda ver que estamos solucionando su problema y que por lo tanto estamos trabajando.

Seguridad: Asegurar y proteger los datos del cliente

Tema: Bases de Datos y Repositorio de datos

5.

DW: es un gran repositorio de datos que almacena gran cantidad de información que proviene de otras fuentes, que servirán para su análisis de datos

DM: este pertenece al DW, que es más simple ya que solo recolecta información mas específica y enfocado a un único, ejemplo , marketing, finanzas, etc.

Data Lake: este es un repositorio más avanzado que almacena datos que serán utilizados en análisis de datos como machine learning y trabaja sobre diferentes fuentes de información como redes sociales, dispositivos conectados a internet, etc.

6.

Repositorio de Datos: es únicamente un lugar no físico, repositorio que almacena y contiene información o datos

Servidor de Base de Datos: Un servidor de base de datos es un software que gestiona y almacena datos. Por ejemplo SQL Server, MySQL.

7.

Ya puede utilizar IaaS, SaaS y PaaS para empezar esto ya facilita su uso.

Escalabilidad: Ajusta recursos fácilmente, ósea que es dinámico su uso.

Costo: Pagas únicamente por lo que almacenas y usas a diferencia de las demás que pagas por el servicio completo lo uses o no.

Accesibilidad: Accede desde cualquier lugar.

Seguridad: Alta protección de la información y de la infraestructura.

Menos mantenimiento: Debido a que el proveedor es el responsable de darle mantenimiento a toda la infraestructura y es algo super notable ya que no te tienes que preocupar por actualizar o dar mantenimiento al sitio.

8.

Para que una organización considere utilizar Big Data, debe tener un sistema muy grande de información, gran volumen de datos, mucha variedad de datos que se complica de encontrar datos en específico, una alta velocidad de procesamiento ya que este mecanismo es especializado para gran repositorio de datos y por lo tanto soporta y actualiza la información en tiempo real prácticamente, un análisis de datos avanzados y escalabilidad es decir que puede ir en crecimiento de datos sin ningún problema.

Ejemplos:

Volumen de Datos: Amazon, que realiza millones de transacciones todos los días

Variedad de Datos: Facebook que procesa datos como textos, imágenes, videos

Velocidad de procesamiento: Maps, que te da una ruta en tiempo real de trazo, en tiempo real de un sinnúmero de rutas.

Análisis avanzado: Bancos, que detectan a diario fraudes bancarios y posibles robos de dinero

Escalabilidad: Google que gestiona ajusta sus recursos dependiendo de la cantidad de búsquedas que haya por segundo.