Propuesta para el mejoramiento en el proceso de análisis de datos de los siniestros viales de Bogotá en Cal y Mayor

Lina María Lozano Yara, Leidy Johana Mendoza Cepeda, Saulo Germán Ruiz, Victor Ricardo Sanabria, Martín Eduardo Flores

Especialización Big Data, Fundación Universitaria Compensar

Proyecto Propuesta de Consultoría

Docente Tutor: Wilson Hernando Soto Urrea

19 de mayo de 2023

Esta obra está bajo una Licencia Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional.

Propuesta para el mejoramiento en el proceso de análisis de datos de los siniestros viales de Bogotá en Cal y Mayor

Lina María Lozano Yara, Leidy Johana Mendoza Cepeda, Saulo Germán Ruiz, Victor Ricardo Sanabria, Martín Eduardo Flores

Propuesta de consultoría presentada como requisito parcial para optar al título de:

Especialista Big Data

Fundación Universitaria Unipanamericana – Compensar
Facultad de Ingeniería, Ingeniería de Especialización en Big Data
Bogotá, Colombia
2023

Resumen

Este proyecto se centra en abordar los desafíos enfrentados por Cal y Mayor en su proceso de análisis y visualización de datos de siniestralidad. Actualmente se realiza el proceso ETL, para su posterior análisis que permite visualizar los datos a través de gráficos y generación de mapas de calor, lo cual puede ser una fase que consume mucho tiempo, es así como el objetivo de este proyecto es proponer una solución que aproveche la tecnología de Big Data para reducir el tiempo de análisis y optimizar la toma de decisiones, mediante la implementación de una estructura de procesamiento de datos optimizada y el uso de técnicas avanzadas de programación, la solución propuesta permitirá a Cal y Mayor agilizar su proceso de análisis y mejorar su capacidad para responder de manera oportuna a los clientes.

Palabras Clave: ETL, Big Data, Cal y Mayor, Siniestralidad, Mapas de Calor, Análisis de Datos

Abstract

This project focuses on addressing the challenges faced by Cal y Mayor in its process of analyzing and visualizing accident data. Currently the ETL process is carried out, for its subsequent analysis that allows to visualize the data through graphs and generation of heat maps, which can be a phase that consumes a lot of time, this is how the objective of this project is to propose a solution that takes advantage of Big Data technology to reduce analysis time and optimize decision making. , by implementing an optimized data processing structure and using advanced programming techniques, the proposed solution will enable Cal y Mayor to streamline its analysis process and improve its ability to respond in a timely manner to customers.

Keywords: ETL, Big Data, Cal y Mayor, Accident data, Heatmap, Data Analyst

CONTENIDO

Resumen 3

Abstrac 4

INTRODUCCCIÓN 9

Antecedentes y Justificación 10

Antecedentes 10

Justificación 11

Formulación del problema 13

Objetivo General 13

Objetivos Específicos 13

Alcances y Limitaciones 14

Alcances: 14

Limitaciones: 14

Marco Teórico 15

Big Data 15

ETL 15

Software 16

Sistemas de información 16

Marco Conceptual 17

Metodología 17

Siniestro Vial 18

Mapas de Calor 18

Gráficas 18

Análisis de datos 18

Información 19

Diseño Metodológico 20

Enfoque de la investigación: 20

Diseño de la investigación 20

Población y muestra 20

Instrumentos de recolección de datos 20

Encuesta: 20

Ficha registro de tiempos: 21

Procedimiento de recolección de datos 21

Encuesta: 21

Ficha registro de tiempo: 21

Procedimiento de análisis de datos 21

Metodología Agile 22

Resultados 24

Evaluación del proceso actual sobre el análisis de datos de siniestralidad. 24

Discusión 28

Conclusiones y Recomendaciones 29

Conclusiones 29

Recomendaciones 30

Bibliografía 31

Lista de Gráficas

Gráfica 1 Resultado de Herramienta Implementada 24

Gráfica 2 Herramienta Implementada 24

Gráfica 3 Conocimiento de Herramientas 25

Gráfica 4 Nivel de Dominio 25

Gráfica 5 Estructura de Actividades del proceso actual 26

Gráfica 6 Matriz de Decisiones 26

Gráfica 7 Comparativos 27

INTRODUCCCIÓN

Hoy en día los datos e información se han convertido en algo esencial para la toma de decisiones, por ello para la empresa Cal y Mayor, experta en consultoría, presenta una alta demanda de respuesta y cumplimiento, la cual se destaca por ser una empresa que tiene como uno de sus objetivos principales contribuir a la disminución de lesiones y muertes por siniestros viales en la población, destacando este tipo de análisis de siniestralidad en los estudios donde es partícipe con el fin de proponer y recomendar medidas de mitigación que favorezcan a la buena toma de decisiones en mejora de la seguridad vial para los diferentes actores.

En el presente documento se quiere proponer una optimización en los procesos de análisis de datos, que permita convertir esos datos en información verídica y de una manera más ágil, de esta forma se plantea una solución mediante nuevos y mejores procesos modelados en el aseguramiento de un paso a paso en el tratamiento de los mismos, con el fin no solo de disminuir los tiempos, si no dé también obtener una mayor fiabilidad, lo anterior se logra por medio de la comprensión de la Big Data.

Este proyecto tendrá uso de herramientas tecnológicas para el ETL y toma de decisiones, se debe tener en cuenta que la propuesta se realizara para la sede de la ciudad de Bogotá, tomando en cuenta que la cantidad de datos que se manejan en la empresa son referente al histórico de los últimos cinco años, donde las variables y restricciones son definidas para este, por ende, el documento puede ser utilizado como guía, y no estrictamente como se estipule.

Antecedentes y Justificación

Antecedentes

En el artículo denominado Conociendo Big Data realizado por Juan Jose Camargo, Jonathan Felipe Camargo y Luis Joyanes de la universidad pedagógica y tecnológica de Colombia tuvo como objetivo estudiar y analizar de manera exhaustiva el estado del arte del Big data en donde se aborda brevemente toda la información básica de Big data, alcances, sus tecnologías y aprovechamiento y se concluyó que las características más importantes en la gestión de Big Data son necesarias para convertir datos no estructurados en información y conocimiento que beneficie tanto a personas como a empresas en la toma de decisiones de forma que pueda servir de guía a quienes deseen sacarle un mayor usufructo a los datos y convertirlos en conocimiento, que les sea útil a las empresas u organizaciones, buscando mayor beneficio en estrategias empresariales. Los teóricos que tuvieron como base fueron B. Hopkins, Beyond the Hype of Big Data. (Camargo Vega, Camargo Ortega, & Jayones Aguilar , 2015)

En el caso del repositorio o trabajo de grado por Montero Fernando, denominado Análisis de datos de accidentalidad vial de la ciudad de Bogotá a partir de datos abiertos y datos obtenidos desde redes sociales, se centra en analizar datos de siniestros viales en la ciudad de Bogotá, con el fin de identificar las variables con mayor tasa de accidentalidad, y entender las razones de las mismas a través de los datos, utilizando técnicas big data, que involucran herramientas tecnológicas de análisis. (Montero, 2022)

Justificación

> Desde los años 60, cuándo se hizo evidente la necesidad de analizar y procesar datos de una manera más ágil y sencilla, han venido naciendo herramientas tecnológicas que facilitan el trabajo, y es bien decir que una de las herramientas más utilizadas han sido las hojas de cálculo, las cuales se popularizaron en los años 80´s, y que a partir de allí se han venido presentando mejoras en la herramienta, así como la implementación que realizó la empresa Lotus Development Corporation, quienes realizaron la incorporación por primera vez de gráficos y operaciones más complejas para mostrar datos, es así como hoy día, la empresa Cal y Mayor, puede permitirse tener como principal herramienta de análisis el software Excel. (Fernández, 2017)

Para la empresa, (Vadafone España, 2018) en la actualidad existen otro tipo de herramientas más robustas con las cuales se puede lograr un análisis de datos más completo en menor tiempo, como por ejemplo Python, R, SPSS, Power BI, entre otras, de las cuales se pueden implementar una o varias, dependiendo la necesidad que posea la empresa, el tipo de análisis que desea realizar y los tipos de datos a tratar.

Ahora bien, teniendo en cuenta que la empresa, en este caso Cal y Mayor, es una empresa dedicada a las consultorías en donde se brindan soluciones respecto a la siniestralidad, podemos decir que la implementación de nuevas herramientas y formas de tratar los datos puede acortar el tiempo de respuesta de los análisis realizados, que posteriormente le permitirá proporcionar soluciones y recomendaciones más precisas basadas en el estudio de los datos, a sus clientes, dado que la competencia en el mercado puede marcar la diferencia cuando se usan herramientas que aportan a las soluciones más inmediatas, hay que tener presente que dentro de la analítica ya

existen estudios sobre el uso de la tecnología para estudiar datos y analizarlos más a profundidad. (Cal y Mayor, s.f.)

Por lo tanto, podemos decir que adoptando o proponiendo una herramienta más robusta para el tratamiento y análisis de datos, se podrían reducir los tiempos en la consolidación de la información para presentarla a los clientes, entonces a la hora de escoger una de las herramientas a utilizar, es necesario saber cuál es la cantidad de datos a tratar, el tipo de análisis, proyecto, u objetivo que se desea alcanzar.

Formulación del problema

La empresa Cal y Mayor carece de la implementación de nuevas herramientas tecnológicas para el análisis de siniestralidad que le permitan realizar un análisis de datos con un menor tiempo de ejecución.

Objetivo General

Proponer la optimización en el proceso para el análisis de datos de la empresa Cal y Mayor, en siniestros viales de Bogotá en los últimos 5 años

Objetivos Específicos

- ✓ Evaluar el proceso actual sobre el análisis de datos de siniestralidad
- ✓ Proponer la herramienta tecnológica adecuada para aumentar la eficiencia de la consolidación y análisis de los informes que se generan.
- ✓ Validar la efectividad del nuevo enfoque mediante pruebas y comparaciones con el proceso actual, demostrando la reducción significativa de tiempo en la consolidación y análisis de los informes generados, así como mejoras en la calidad y precisión de los resultados.

Alcances y Limitaciones

El presente documento se realizará bajo los siguientes parámetros de alcances y limitaciones así

Alcances:

- ✓ Se trabajará con la empresa Cal y Mayor, utilizando su información y datos existentes relacionados con la siniestralidad vial.
- ✓ Se implementarán herramientas tecnológicas adecuadas para agilizar y optimizar el flujo de trabajo de análisis de datos.
- ✓ La herramienta tecnológica propuesta, estará dirigida a mejorar el proceso de extracción, transformación y generación de informes de siniestralidad vial en la empresa.
- ✓ Se buscará validar y demostrar la efectividad del nuevo enfoque mediante pruebas y comparaciones con el proceso actual.

Limitaciones:

- ✓ El proyecto estará limitado al análisis de datos de siniestralidad vial únicamente en la ciudad de Bogotá.
- ✓ Se trabajará exclusivamente con la empresa Cal y Mayor, por lo que los resultados y conclusiones obtenidos serán aplicables dentro de su contexto y alcance.
- ✓ Las mejoras propuestas estarán basadas en la disponibilidad y calidad de los datos existentes proporcionados por la empresa.
- ✓ La implementación de las mejoras estará sujeta a los recursos tecnológicos y humanos disponibles en la empresa.
- ✓ El impacto y los resultados obtenidos se evaluarán en función de la eficiencia y la calidad del proceso de análisis de datos de siniestralidad vial, sin abordar otros aspectos relacionados con la seguridad vial o la toma de decisiones en general.

Marco Teórico

Para iniciar, se presentan algunas definiciones sobre el término Big Data, del cual existen innumerables definiciones, entre ellas se tienen:

Big Data

Algunos autores, afirman lo siguiente:

> Forrester define el Big Data es definido como las técnicas y tecnologías que hacen que sea económico hacer frente a los datos a una escala extrema, el Big Data trata de tres cosas: 1) Las técnicas y la tecnología, lo que significa que la empresa tenga personal, el cual tenga gran representación y análisis de datos para tener un valor agregado con información que no ha sido manejada. 2) Escala extrema de datos que supera a la tecnología actual debido a su volumen, velocidad y variedad. 3) El valor económico, haciendo que las soluciones sean asequibles y ayuden a la inversión de los negocios". (Hopkins, 2011)

Según (Kusnetzky, 2010), Big Data "se refiere a las herramientas, los procesos y procedimientos que permitan a una organización crear, manipular y gestionar conjuntos de datos muy grandes y las instalaciones de almacenamiento".

Para (Salgado, 2014) el Big Data es definido como "un gran volumen, velocidad o variedad de información que demanda formas costeables e innovadoras de procesamiento de información que permitan ideas extendidas, toma de decisiones y automatización del proceso.

ETL

De acuerdo a la publicación, en la plataforma de los IEBS, Los procesos ETL (Extract, Transform, Load) hacen referencia a un conjunto de técnicas, herramientas y tecnologías

> que permiten extraer datos de varias fuentes, transformarlos de forma que sean veraces y útiles, y cargarlos en otros sistemas con el fin de que puedan ser accesibles por los niveles de la organización que lo requieran. (Bello, 2022)

La construcción del ETL puede dividirse en tres subprocesos o componentes: componente de extracción, componente de transformación y componente de carga.

Software

> Según se describe en el libro de Ingeniería de Software, “no son solo programas, sino todos los documentos asociados y la configuración de datos que se necesitan para hacer que estos programas operen de manera correcta.” Es decir, el software en si es un compendio de actividades previas y posteriores a la creación de un programa que tienen una estructura definida y se debe seguir un proceso adecuado desde su análisis hasta la finalización con la creación de un producto de software que al final debe llevar una documentación orientada a el usuario final indicando como se debe utilizar, al igual se entrega con toda la información generada durante el proceso de desarrollo. (Sommerville , 2011)

Sistemas de información

> Dentro de este marco, se debe hablar del impacto que ha tenido los sistemas de información en las organizaciones, desde la automatización de procesos. Esto ha permitido a las empresas mejorar sus procesos haciéndolos más agiles y competitivas frente a un mercado cada vez más global. Desde esta perspectiva se disponen las siguientes temáticas.

Las Tecnologías de la Información y la Comunicación (TIC) permiten a las empresas mejorar su eficiencia mediante, por ejemplo, la reducción de costes o la mejora de la calidad y del servicio ofrecido a los consumidores. Por tanto, la adopción de estas tecnologías resultará fundamental para la competitividad de las pequeñas medianas empresas (MI PYMES) tanto por las oportunidades que ofrece como por las amenazas que se derivan para su supervivencia si no se adaptan a los cambios introducidos por estas tecnologías en los mercados (Gargallo y Ramírez,2007. p 370).

Los sistemas de información han hecho que las organizaciones sean más competitivas y generadoras de valor constantemente como lo indica Tobos (2010) quien indica que "Cualquier empresa que implementa SIG persigue la excelencia operativa, eliminar distancias, conocer nuevos productos y modelos de negocios, buenas relaciones con los clientes y los proveedores, toma mejorada (racional) de decisiones, ventajas competitivas y supervivencia en un mundo globalizado" (p. 6)

Marco Conceptual

Metodología

Como metodología se denomina la serie de métodos y técnicas de rigor científico que se aplican sistemáticamente durante un proceso de investigación para alcanzar un resultado teóricamente válido. En este sentido, la metodología funciona como el soporte conceptual que rige la manera en que aplicamos los procedimientos en una investigación. (Significados, 2019)

Siniestro Vial

Según (Monroy, 2021)Un siniestro vial es aquel acontecimiento que cuenta con imprevistos donde se ve involucrado cualquier tipo de vehículo ya sea un automóvil, motocicleta, camión, etc. Al ocurrir este tipo de situaciones suelen dejar daños tanto materiales como personales**.**

Mapas de Calor

> Según el glosario Digital Arimetrics, los mapas de calor son una representación gráfica de los datos utilizando colores para indicar el nivel de actividad. Por lo general, se emplean colores más oscuros para indicar una baja actividad y colores más brillantes para indicar una alta actividad. (Arimetrics, s.f.)

Gráficas

Según (Westreicher, s.f.) El gráfico es una herramienta para representar una serie de datos por medio de un instrumento visual. De esta forma, se intenta ilustrar, entre otros, la relación entre variables estadísticas o la evolución de estas en el tiempo.

Análisis de datos

> El análisis de datos es la ciencia que se encarga de examinar un conjunto de datos con el propósito de sacar conclusiones sobre la información para poder tomar decisiones, o simplemente ampliar los conocimientos sobre diversos temas.
>
> El análisis de datos consiste en someter los datos a la realización de operaciones, esto se hace con la finalidad de obtener conclusiones precisas que nos ayudarán a alcanzar

nuestros objetivos, dichas operaciones no pueden definirse previamente ya que la recolección de datos puede revelar ciertas dificultades. (QuestionPro, s.f.)

Información

Como información denominamos al conjunto de datos, ya procesados y ordenados para su comprensión, que aportan nuevos conocimientos a un individuo o sistema sobre un asunto, materia, fenómeno o ente determinado. La palabra, como tal, proviene del latín informatĭo, informatiōnis, que significa ‘acción y efecto de informar” (Significados, 2023)

Diseño Metodológico

Enfoque de la investigación:

El enfoque escogido para esta propuesta ha sido cuantitativo, bajo un proceso secuencial, teniendo en cuenta que se desea tener información de la cantidad de personas que poseen conocimiento de una herramienta para análisis de datos diferente a la que maneja actualmente en la empresa e identificar cual es esa herramienta que tendría mayor aceptación y adaptación.

También debido a que se quiere analizar los tiempos de respuesta del software usado actualmente en la empresa y el que se desea proponer, realizando un comparativo de tiempo de las dos herramientas de análisis de datos.

Diseño de la investigación

El diseño dentro de nuestra propuesta será no experimental de tipo descriptivo

Población y muestra

La población y muestra escogida para este proyecto, son los empleados de la empresa Cal y Mayor, de la sede en Colombia Bogotá

Instrumentos de recolección de datos

Encuesta:

Se selecciono la encuesta como instrumento de recolección de datos, que nos permita saber, si actualmente dentro de la empresa los empleados han tenido alguna experiencia con

herramientas de análisis de datos diferentes a Excel, y su grado de conocimiento de estas en caso de que aplique, esto con el fin de determinar cuál sería la herramienta más adecuada que se propondría para la implementación, que no genere una inversión significativa de recursos.

Ficha registro de tiempos:

Se recolectarán datos a través de un registro de tiempos por tarea realizada en una ficha que nos permita visualizar y comparar la herramienta a proponer y la que se usa actualmente en la empresa para el análisis de datos.

Procedimiento de recolección de datos

Encuesta:

Se aplicará la encuesta digitalmente utilizando formulario en línea que permita al participante o encuestado contestar desde sus dispositivos electrónicos.

Ficha registro de tiempo:

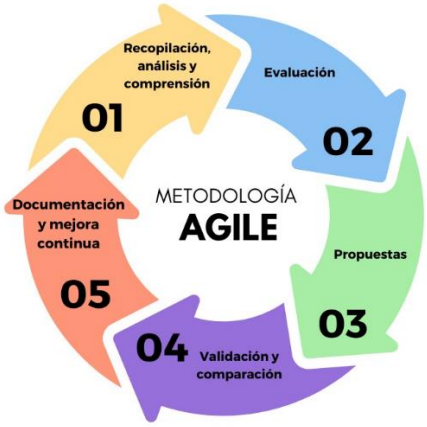
El procedimiento en esta recolección será de manera manual, identificando primero los dos tipos de herramientas a evaluar y las tareas específicas a realizar en cada una, para así poder registrar los tiempos de ejecución, con el fin de poder analizar los datos de tiempo arrojado.

Procedimiento de análisis de datos

El análisis de datos para este proyecto se concentrará en la cantidad de personas que tienen conocimiento en el manejo de otra herramienta para análisis de datos y la herramienta más conocida por los mismos

Y el segundo análisis se basa en identificar, cual, de las herramientas a calificar, es óptima en tiempos de respuesta para el análisis de los datos de siniestralidad de la empresa.

Metodología Agile



1. Recopilación de datos: Obtener los datos sobre las herramientas que se utilizan dentro de la empresa y la que sería más fácil y optima de proponer según el conocimiento de los colaboradores de Cal y Mayor
2. Análisis y comprensión de los datos: Realizar una exploración inicial de los datos para comprender su estructura, calidad y posibles problemas.
3. Evaluación del proceso actual: Evaluar el proceso actual de análisis de datos de siniestralidad vial en Cal y Mayor. Identificar fortalezas, limitaciones y áreas de mejora.
4. Propuesta de nueva herramienta: Basado en la evaluación del proceso actual y en los objetivos del proyecto, formular una nueva herramienta tecnológica para el análisis de datos de siniestralidad vial. Esto implica definir las etapas del proceso, las técnicas y herramientas a utilizar, y los criterios de evaluación y validación.

5. Validación y comparación: Validar la efectividad de la herramienta propuesta mediante pruebas y comparaciones con el proceso actual, para poder evaluar la eficiencia, calidad y precisión de los resultados generados,
6. Documentación y Mejora continua: Documentar todo el proceso metodológico, incluyendo los resultados obtenidos y las conclusiones, que permita realizar ajustes, mejoras en la herramienta, en el proceso de análisis con base en los resultados y la validación.

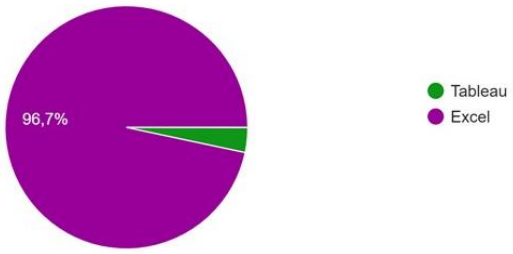
Este diseño metodológico proporciona una guía paso a paso para abordar el proyecto, desde la recopilación de datos hasta la implementación y evaluación del nuevo enfoque en el análisis de datos de siniestralidad vial. Es importante adaptar el diseño a las necesidades y recursos específicos de la empresa y asegurarse de mantener un enfoque sistemático y riguroso a lo largo del proceso

Resultados

Evaluación del proceso actual sobre el análisis de datos de siniestralidad.

Recopilación y Análisis de datos: Dentro del desarrollo, se realizó una encuesta a los colaboradores de la empresa de la sede de Bogotá que nos arrojó los siguientes resultados.

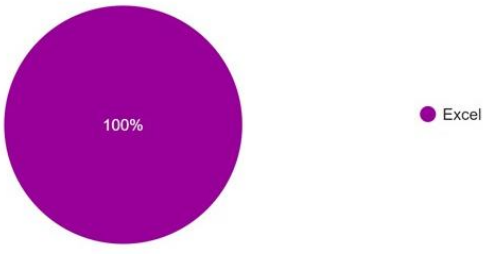
Gráfica 1 Resultado de Herramienta Implementada



Fuente: Elaboración propia

Se puede evidenciar que en la empresa Cal y Mayor, la herramienta mas usada es Excel, con un porcentaje de 96,7% y Tableau con un 3.3%

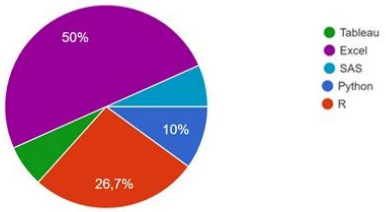
Gráfica 2 Herramienta Implementada



Fuente: Fuente: Elaboración propia

El 100% del trabajo de análisis realizado en la empresa Cal y Mayor se realiza con Excel

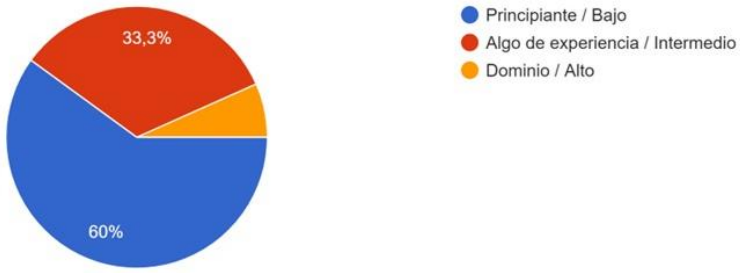
Gráfica 3 Conocimiento de Herramientas



Fuente: Fuente: Elaboración propia

De las opciones dadas, sobre el conocimiento de las herramientas para análisis de datos, tales como Tableau, Excel, SAS, Python, y R, se puede evidenciar que el 50% de los empleados sólo posee conocimientos en Excel, seguido de un 26,7% quienes respondieron tener conocimiento sobre R y el 10% sobre Python.

Gráfica 4 Nivel de Dominio



Fuente: Elaboración propia

Podemos observar que el 60% de los encuestados tiene un nivel de expertiz principiante o bajo, el 33,3% cuenta con algo de experiencia. y sólo el 6,7% un dominio alto.

Gráfica 5 Estructura de Actividades del proceso actual

NO	Detalle de la actividad	Tiempo (h)	%
1	Extracción de la información general (Base de datos de siniestros y actor vial)	2	5%
2	Filtro por zona de influencia	5	12%
3	Análisis y procesamiento de información (Asignación de factores según gravedad de siniestro)	8	20%
4	Elaboración de gráficos de variables cualitativas y cuantitativas	7	17%
5	Generación de mapas de calor para identificar puntos críticos	14	34%
6	Redacción de informe situación actual	5	12%
	Total	41	100%

Fuente: Elaboración propia

La anterior es una ficha de actividades que refleja el tiempo de ejecución de análisis de datos actualmente en la herramienta de Excel.

Evaluación del proceso: De acuerdo a los datos arrojados en la encuesta, se procede a realizar matriz de decisión para elegir la mejor herramienta a implementar.

Gráfica 6Matriz de Decisiones

Característica	Peso	Python	R
Entornos de trabajo	2	3x2=6	2x2=4
Analisis de Datos	3	3x3=9	2x3=6
Visualizaciónde datos	4	3x4=12	2x4=8
Transformación de datos	2	2x2=4	2x2=4
Paqetes y librerias	2	3x2=6	2x2=4
Total		37	26

Fuente: Elaboración Propia

Propuesta de nueva Herramienta: De acuerdo a la evaluación y ponderación en el proceso para la identificación de la tecnología a proponer, por medio de la matriz de decisiones, y teniendo en cuenta las características de cada una, se puede decir que, bajo los criterios expuestos, en aplicación y funcionalidad, la mejor herramienta es PYTHON.

Validación y comparación: Teniendo en cuenta el punto anterior, se procedió a realizar pruebas de ejecución de análisis de datos, con Datasets de siniestralidad, en la cual se buscaba medir el tiempo con Excel y el tiempo de ejecución con Python.

Gráfica 7 Comparativos

NO	Detalle de la actividad	Tiempo (h) Excel	%	Tiempo (h) Python	%
1	Extracción de la información general (Base de datos de siniestros y actor vial)	2	5%	2	15%
2	Filtro por zona de influencia	5	12%	5	37%
3	Análisis y procesamiento de información (Asignación de factores según gravedad de siniestro)	8	20%	0,25	2%
4	Elaboración de gráficos de variables cualitativas y cuantitativas	7	17%	0,5	4%
5	Generación de mapas de calor para identificar puntos críticos	14	34%	0,75	6%
6	Redacción de informe situación actual	5	12%	5	37%
	Total	41	100%	13,5	100%

Fuente: Elaboración propia

Discusión

- ✓ Teniendo en cuenta las pruebas realizadas para valorar el tiempo de ejecución en el análisis de datos, con la herramienta actual y la propuesta, es necesario aclarar, que, si la empresa realiza la implementación de la tecnología presentada, el tiempo de ejecución tiende a disminuir dado que para el proceso sólo se necesita codificar una vez, porque el proceso ya se vuelve automático una vez cargados los datos.
- ✓ La propuesta que se esta realizando por medio de esta consultoría, a la empresa Cal y Mayor, beneficiaria sus resultados, teniendo en cuenta que estamos en la era de la digitalización y generación de grandes volúmenes de datos, lo que le permite generar mas competitividad en el mercado.

Conclusiones y Recomendaciones

Conclusiones

- ✓ La revisión y comprensión de un proceso de primera mano y por actores ajenos a los que están relacionados con el proceso permite identificar aspectos a mejorar y cuales aspectos se pueden optimizar a través del aprovechamiento de las tecnologías adecuadas y con asesoramiento profesional y calificado.
- ✓ De acuerdo al análisis de un proceso relacionado con obtención de datos es posible establecer una optimización y mayor uso de los recursos disponibles a través de la aplicación de Big Data y técnicas como ETL que permiten ser base para la toma de decisiones encaminadas al fortalecimiento de una organización
- ✓ En la formación de la especialización en Big Data se realiza un acercamiento a herramientas específicas que permiten de acuerdo a las necesidades y escenarios, mostrar opciones para la correcta gestión, interpretación y presentación de los datos de un proceso o de varios procesos que se quieran mejorar al interior de una organización.

Recomendaciones

- ✓ Capacitación continua a los colaboradores de la empresa, para el manejo de una nueva herramienta de análisis de datos
- ✓ una vez que las herramientas de automatización estén implementadas, es importante contar con un plan de mantenimiento y actualización a mediano o largo plazo. Esto incluye la corrección de errores, mejoras en el rendimiento, actualización de bibliotecas utilizadas y la incorporación de nuevas funcionalidades que se ajusten a las nuevas necesidades.

Bibliografía

Arimetrics. (s.f.). *Glosario Digital Arimetrics*. Obtenido de Mapas de Calor: https://www.arimetrics.com/glosario-digital/mapa-de-calor

Bello, E. (7 de Enero de 2022). *IEB School*. Obtenido de Guía de Procesos ETL: Qué son, cómo usarlos y herramientas clave: https://www.iebschool.com/blog/que-son-los-procesos-etl-big-data/#:~:text=Los%20procesos%20ETL%20(Extract%2C%20Transform,niveles%20de%20la%20organizaci%C3%B3n%20que

Cal y Mayor. (s.f.). *Cal y Mayor*. Obtenido de Quienes Somos: https://www.calymayor.com.mx/mx/quienes_somos.html

Camargo Vega, J. J., Camargo Ortega, J. F., & Jayones Aguilar , L. (2015). Conociendo Big Data. *Revista Facultad de Ingeniería Universida Pedagogica y Tecnológica de Colombia*, 63-77.

Fernández, Y. (6 de Octubre de 2017). *Xataka*. Obtenido de Historia de las hojas de cálculo Digitales: https://www.xataka.com/historia-tecnologica/la-historia-de-las-hojas-de-calculo-digitales-de-idea-descartada-a-herramienta-imprescindible

Gallardo Castel, A., & Ramirez Alesón , M. (Septiembre de 2007). *Redalyc*. Obtenido de La adopción de las Tecnologías de la Información en pequeñas y medianas Empresas: https://www.redalyc.org/articulo.oa?id=477748626003

Hopkins, B. (28 de Octubre de 2011). *CIO*. Obtenido de Beyond the Hype of Big Data: https://www.cio.com/article/284795/data-management-beyond-the-hype-of-big-data.html

Kusnetzky, D. (15 de Febrero de 2010). *ZDNet*. Obtenido de What is "Big Data?": https://www.zdnet.com/article/what-is-big-data/

Monroy, J. (29 de Noviembre de 2021). *Migo Seguros*. Obtenido de Siniestro Vial: https://migoseguros.com/articulos/siniestro-vial/

Montero, F. V. (18 de Marzo de 2022). *Análisis de datos de Accidentalidad en la ciudad de Bogotá, a partir de datos abiertos y obtenidos de redes sociales*. Obtenido de Universidad Nacional de Colombia: https://repositorio.unal.edu.co/bitstream/handle/unal/81571/1032474552.2022.pdf?sequence=1&isAllowed=y

QuestionPro. (s.f.). *QuestionPro*. Obtenido de Análisis de Datos: https://www.questionpro.com/es/analisis-de-datos.html#que_es_analisis_de_datos

Salgado, M. (31 de Enero de 2014). *Computer World*. Obtenido de Oracle apuesta por Big Data con tecnología y proyectos: https://www.computerworld.es/big-data/oracle-apuesta-por-big-data-con-tecnologia-y-proyectos

Significados. (17 de Mayo de 2019). *Significados*. Obtenido de Definición de Metodología: https://www.significados.com/metodologia/

Significados. (19 de Mayo de 2023). *Significados*. Obtenido de Inofrmación Definición : https://www.significados.com/informacion/

Sommerville , I. (2011). Ingeniería de Software. En S. Ian, *Ingeniería de Software* (pág. 5). México: Pearson Education .

Tobos Montilla, J. C. (17 de Agosto de 2010). *Univer*. Obtenido de La Importancia de los Sistemas de Información Gerencial en la toma de Decisiones: https://repository.unimilitar.edu.co/bitstream/handle/10654/3697/TobosMontillaJuanCarlos2010.pdf%20%20%20%2026-09-2017;jsessionid=0F80CA820BF39C900F7E3C5E08AE9A51?sequence=2

Vadafone España. (23 de Enero de 2018). *Vadafone*. Obtenido de Big Data: desde los inicios hasta hoy: https://ideasparatuempresa.vodafone.es/big-data-desde-los-inicios-hoy/

Westreicher, G. (s.f.). *Economipedía*. Obtenido de Gráfico: https://economipedia.com/definiciones/matematicas/grafico.html