

ANALÍTICA DE DATOS EN ARCGIS INSIGHTS

ELIANA ALEJANDRA MUÑOZ FAJARDO

UNIVERSIDAD SANTO TOMÁS SECCIONAL TUNJA

FACULTAD DE INGENIERÍA ELECTRÓNICA

TUNJA

2023

PRÁCTICA PROFESIONAL “ANÁLITICA DE DATOS EN ARCGIS INSIGHTS”

ELIANA ALEJANDRA MUÑOZ FAJARDO

Pasantía realizada como opción de grado para optar el título de Ingeniero Electrónico

Director: Mg. Ing. Pablo Andrés Álvarez Camargo

UNIVERSIDAD SANTO TOMÁS SECCIONAL TUNJA

FACULTAD DE INGENIERÍA ELECTRÓNICA

TUNJA

2023

Exoneración de Responsabilidades

Como estudiante de la universidad Santo Tomas Entiendo que soy responsable de todos los conceptos, investigaciones, actividades, análisis y conclusiones involucrados con mi participación en las actividades fuera del campus de la Universidad. Asumo completa responsabilidad de todo lo estipulado dentro de este proyecto, eximo a la facultad de ingeniería electrónica y por supuesto a la universidad de cualquier responsabilidad.



Eliana Alejandra Muñoz Fajardo

Nota de aceptación

Aceptada como opción de grado, dado que

Ha cumplido con los objetivos propuestos y

Los requisitos exigidos por la facultad de

Ingeniería Electrónica

Ing. Pablo Andrés Álvarez Camargo (Docente Tutor)

Firma del presidente del Jurado

Firma del Jurado

Firma del Jurado

DEDICATORIA

En primer lugar, dedico este trabajo a Dios por darme salud y vida para llegar a este punto de la carrera, por cada una de las bendiciones que me ha regalado para superar los retos que se me han presentado, por las fortalezas y habilidades que me ha brindado día a día para seguir adelante. A mis padres *Irma Mariela Fajardo González* y *Aristóbulo Muñoz Pardo* un gran ejemplo a seguir, por el apoyo incondicional a lo largo de mi vida, por cada una de las enseñanzas generadas hasta el día de hoy, por la paciencia y el amor que han sido motivo en mi crecimiento tanto personal como profesional.

En segundo lugar, a mis hermanos, por cada uno de los consejos dados, por el apoyo y la colaboración para superar cada uno de los inconvenientes presentados a lo largo de mi formación universitaria, al ingeniero *David Santiago González Torres*, por la motivación para continuar cuando ya no quería más, a todos mis amigos por la paciencia y el apoyo, por participar y acompañarme en este camino.

Finalmente dedico este trabajo a todas aquellas personas que hicieron parte de mi formación académica y profesional en los últimos 5 años de carrera universitaria, por las experiencias vividas y por los conocimientos compartidos.

AGRADECIMIENTOS

Primeramente, quiero agradecerle a Dios por permitirme alcanzar cada uno de los logros que he tenido durante toda mi vida, por la vida y la salud, por las bendiciones que nos ha brindado siempre a mi familia y a mí. A mis padres *Irma Mariela Fajardo González* y *Aristóbulo Muñoz Pardo* por todo lo que me han dado, por el amor y apoyo incondicional en todo momento, por creer siempre en mis capacidades y habilidades, por forjarme cada día más para ser la mejor, por dar sin esperar nada a cambio, por los consejos y las enseñanzas, por el ejemplo, la orientación y la dedicación, a ellos les agradezco todo lo que soy hoy.

En segundo lugar al Ministerio de Trabajo por suministrar convocatorias para realizar la práctica laboral en las diferentes entidades estatales del país, a la Superintendencia de Transporte por abrirme las puertas y brindarme la oportunidad de ingresar al campo profesional de la manera más amable y enriquecedora, al arquitecto *Esteban Martínez Torres* y al ingeniero *Oscar Santiago Acevedo*, por la confianza y seguridad que me dieron para fortalecer mucho más mis conocimientos e inculcarme unos nuevos, al ingeniero *Pablo Andrés Álvarez Camargo*, por orientarme en mi formación técnica, sin su apoyo no hubiese sido posible alcanzar este logro tan importante.

Por último, a la facultad de ingeniería electrónica de la Universidad Santo Tomás, por brindarme el espacio para recibir preparación académica de alta calidad, a sus docentes quienes me formaron para ser profesional y a todo el cuerpo administrativo de la institución. A mis compañeros de carrera, quienes me guiaron cuando no todo era claro y a quienes trabajaron conmigo por años.

Tabla de Contenido

1. Resumen.....	11
2. Abstract.....	12
3. Introducción	13
4. Planteamiento del problema	15
5. Justificación	16
6. Objetivos	18
6.1. General.....	18
6.2. Específicos.....	18
7. Marco teórico	19
8. Estado del arte.....	22
9. Desarrollo del proyecto.....	25
9.1. Diagrama de Flujo	25
9.2. Planificación y diseño inicial	26
9.3. Creación del Formulario	27
9.4. Publicación y recopilación de datos.....	32
9.5. Almacenamiento y preparación de datos para ArcGIS Insights	36
10. Análisis de Resultados y Discusión	38
10.1. Comparativo ArcGIS Insights, Power Bi y Tableau.....	38
10.1.1. Power Bi Desktop	38
10.1.2. Tableau Desktop.....	40
10.1.3. Fortalezas y debilidades entre ArcGis Insights, Power Bi y Tableau	41
10.2. Análisis de datos ArcGIS Insights	44
10.2.1. Datos por fecha.....	46
10.2.2. Datos por nivel de gravedad.....	53
10.2.3. Tendencia en los datos	58
10.2.4. Informes de presentación	63
11. Conclusiones y Trabajos futuros	66
12. Glosario	68
13. Referencias Bibliográficas	70
14. Anexos	72

14.1.	Presupuesto.....	72
14.2.	Estrategias de divulgación.....	74

Lista de Figuras

Figura 1. Diagrama de Flujo de trabajo	25
Figura 2. Back-end del formulario PESCRI basada en Microsoft Excel.	28
Figura 3. Front-end del formulario en Survey123 Connect antes de publicarlo	30
Figura 4. Función usada para extraer los datos del archivo operadores	31
Figura 5. Programación de función en Java Script	31
Figura 6. Aplicativo ArcGIS Survey123 para el diligenciamiento de la información.....	32
Figura 7. Interfaz ingreso al Formulario	33
Figura 8. Prueba de encuesta PESCRI.....	34
Figura 9. Información sobre el siniestro.....	35
Figura 10. Base de datos PESCRI	36
Figura 11. Tablero de control PESCRI	37
Figura 12. Interfaz gráfica datos PESCRI en Power Bi.....	39
Figura 13. Interfaz gráfica datos PESCRI en Tableau.	40
Figura 14. Interfaz gráfica PESCRI en ArcGIS Insights.....	44
Figura 15. Ventana Mapas.....	45
Figura 16. Ponderación por color.....	46
Figura 17. Mapa de ubicación de los siniestros ocurridos en el año 2023	47
Figura 18. Diagrama de cuerdas día de la semana respecto a la condición climática	48
Figura 19. Relación y correlación entre Día/gravedad y Día/condición climática	49
Figura 20. Mapa de calor horario y día del siniestro.....	50
Figura 21. Diagrama de barras siniestros por mes.....	51
Figura 22. Análisis Relación y correlación de datos por fecha	52
Figura 23. Gráfico de anillos-Nivel de gravedad.....	53
Figura 24. Diagrama de cuerdas- gravedad vs curva o recta.....	54
Figura 25. Dato de reloj-Gravedad vs condición climática	55
Figura 26. Mapa de árbol-Tipo de siniestro	56
Figura 27. Relación y correlación de los datos por nivel de gravedad	57
Figura 28. Gráfico de barras, ponderación de siniestros por concesión	58
Figura 29. Serie de tiempo siniestros durante el 2023	60
Figura 30. Descripción y estadística de los datos	61
Figura 31. Referencias rápidas para el informe.....	63
Figura 32. Informe reporte de PESCRI estadística	64
Figura 33. Vista general de todos los datos.....	65
Figura 34. Publicidad para evento Gisday	75
Figura 35. Publicación en Flickr por parte de Esri Colombia	76
Figura 36. Evento Gisday con la Supertransporte	76
Figura 37. Ganadores de una de las mejores propuestas futuras con el aplicativo.....	77
Figura 38. Acompañamiento del director de concesiones e infraestructura y todo el equipo de trabajo ..	78

Lista de Tablas

Tabla 1. Presupuesto	73
----------------------------	----

1. Resumen

En este documento se describe el proceso realizado en cada una de las actividades que se desarrollaron durante la práctica laboral como pasante en la Superintendencia de Transporte la cual es una entidad que se destaca a nivel nacional en todo el país, esto durante el primer periodo del 2023, donde se llevaron a cabo labores de ingeniería electrónica apoyando a la Delegatura de Concesiones e Infraestructura en el análisis de datos recolectados en el marco de la supervisión objetiva y subjetiva, así como en el ajuste de las actividades correspondientes a la apropiación tecnológica en los procesos misionales de inspección y vigilancia, determinando por un lado el enfoque y método de recolección de los datos que se quieren obtener y por el otro la fuente a utilizar para el análisis e interpretación de la información. Dentro de la apropiación tecnológica que se maneja para el supervisado de los procesos de vigilancia, se trata puntualmente el software ArcGIS Survey 123 para la recolección de los datos y ArcGIS Insights para el posterior análisis de la información, dando a conocer la manera en que se pueden encontrar relaciones y correlaciones entre los datos y con ello hallar análisis de tendencias de estos en pro de la buena toma de decisiones.

Palabras clave: GIS, ArcGIS Survey123, ArcGIS Insights, correlación, patrones, tendencias.

2. Abstract

This document describes the process carried out in each of the activities that were carried out during the work experience as an intern at the Superintendency of Transportation during the first period of 2023, where electronic engineering work was carried out supporting the Delegation of Concessions and Infrastructure in the analysis of data collected in the framework of objective and subjective supervision, as well as in the adjustment of the activities corresponding to the technological appropriation in the inspection and surveillance missionary processes, determining on the one hand the focus and method of collection of the data to be obtained and, on the other hand, the source to be used for the analysis and interpretation of the information. Within the technological appropriation that is managed for the supervision of the surveillance processes, the ArcGIS Survey 123 software for data collection and ArcGIS Insights for the subsequent analysis of the information are dealt with, disclosing the way in which They can find relationships and correlations between the data and thus find analysis of trends of these in favor of good decision making.

Keywords: GIS, ArcGIS Survey123, ArcGIS Insights, Correlations, Patterns, Trends.

ANALÍTICA DE DATOS EN ARCGIS INSIGHTS

3. Introducción

El manejo de la data es una de las tareas con mayor prioridad en la actualidad, es por ello que a medida que va pasando el tiempo, se van originando nuevas aplicaciones o herramientas de software para suplir cada necesidad; en este caso ArcGIS Insights tiene la capacidad de fusionar el análisis proveniente de la ubicación con flujos de trabajo, dentro de él se pueden llevar a cabo análisis de carácter espacial, estadístico y sobre todo predictivo, lo cual si se entra a ver detalladamente lo que se busca y frente a que necesidades, entonces se direcciona la solución. Si bien se sabe lo importante no es tener la información sino lo que se hace con ella, la verdadera esencia del análisis es poder explotar los datos.

Teniendo en cuenta que la Superintendencia de Transporte tiene como propósito velar por el control, inspección y vigilancia de este sector, cuenta con un gran manejo de datos en donde se presentan diferentes enfoques, es por ello que para cada una de las categorías referentes a los programas especiales, existen diferentes temarios que permiten la recolección de datos para su posterior análisis y finalmente la toma de decisiones y acciones; algunos de los programas que hacen parte de la categoría Aérea, son: ETA'S Tiquetes (Empresas de Transporte Aéreo), PEACCI (Programa Especial de Accesibilidad e Inclusividad), SETA (Supervisión Especial Temporada Alta), en la categoría férrea se encuentra: OF(Operadores Férreos), IFC(Infraestructura férrea concesionada), IFNC(Infraestructura férrea no concesionada) y finalmente dentro de la categoría de carretero se tiene: PESCRI(Programa Especial Sectores Críticos de Siniestralidad), PECSO(Programa Especial Control al Sobrepeso), IP/REV (Interoperabilidad de Recaudo Electrónico Vehicular), SETA Carretero(Supervisión Especial Temporada Alta) y Terminales. Así

mismo es importante aclarar que cada una de ellas manejan un objetivo específico a cumplir; estos programas son el pilar de investigación, son la clave para realizar o no una acción, bien sea de carácter normativo o informativo. Partiendo que dentro de estos programas se desarrollan los formularios con preguntas de relevancia alta e importante, para que al momento de obtener las respuestas se lleve a cabo un análisis tanto espacial como estadístico, y posteriormente tomar una decisión, saber qué es lo que se puede hacer de acuerdo con lo obtenido en los resultados.

A lo largo de este documento se evidenciará en detalle todo el proceso desde la recolección de los datos hasta el tratamiento y análisis de la información obtenida respecto a uno de los programas especiales que anteriormente se mencionó, se habla de PESCRI (programa especial sectores críticos de siniestralidad), lo cual con ayuda de ArcGIS en sus diferentes módulos se logrará obtener la dirección a la que se debe apuntar como entidad responsable en el sector de transporte.

Finalmente con este método de adquisición de datos, es posible obtener un sin número de observaciones que normalmente no se veían a simple vista en los demás tipos de análisis que se manejaban en otras plataformas, este software se caracteriza por manejar un sistema de información geográfica, es decir, permite la visualización de estos en mapas o gráficos interactivos para llevar a cabo un buen entendimiento sobre los distintos patrones espaciales y/o temporales, entre otras cosas que serán clave para poner en marcha un buen tratamiento de datos, el cual viene dado por el análisis estadístico para descubrir relaciones y correlaciones entre las diferentes variables presentes.

4. Planteamiento del problema

Actualmente la Superintendencia de Transporte vela por el cumplimiento de su objetivo el cual abarca la vigilancia, la inspección y el control que regula la prestación del servicio público de transporte, puertos, concesiones e infraestructura, servicios conexos y protección a usuarios de sector de movilidad, todo esto bajo unas condiciones normativas. Con base en lo anterior es importante mencionar las diferentes direcciones que se manejan dentro de la entidad. Por un lado, está la dirección de promoción y prevención y dentro de ella la delegatura de concesiones e infraestructura que es el foco central de la problemática. Dentro de esta delegatura se manejan programas especiales para el cumplimiento del objetivo relacionado con la vigilancia, inspección y control, estos programas especiales están diseñados dependiendo a la categoría en que se requiera, bien sea en la parte aérea, férrea o carretera, para este caso el programa de enfoque viene dado por el sector carretero.

De acuerdo a ello el problema radica en que mensualmente se reciben miles de datos dentro del reporte en cada programa, haciendo claridad que cada categoría contiene cierto número de vigilados y cada vigilado puede tener una o más concesiones, lo que al realizar el análisis de esta información e identificar las fallas que se están presentado, es tarea de cada uno de los encargados de los programas, sin embargo esto tarda horas e incluso días para obtener un informe, ya que en la mayoría de casos se debe hacer un cruce de datos entre lo que reportan los vigilados y lo que reportan otras entidades.

¿Cómo se podría automatizar el proceso de recolección de datos para encontrar relaciones y correlaciones, que permitan establecer tendencias y estrategias de prevención en pro de la toma de decisiones a futuro?

5. Justificación

Existen varias alternativas que a nivel de software se pueden usar, pero ninguna como esta, las demás proporcionan la opción de realizar un análisis empresarial de carácter más general, mientras que ArcGIS Insights es ideal para aquellos que necesitan trabajar con datos geográficos y eso es lo que se pretende implementar en la entidad, un análisis espacial avanzado que permita realizar estudios de ubicación, análisis de patrones, integración de datos geográficos, visualización de mapas interactivos y análisis temporal lo cual permita observar la manera en que cambian los datos con el tiempo.

Este proyecto contiene muchos aspectos potenciales y beneficiosos en su desarrollo para la Superintendencia de transporte, de carácter institucional puede tener gran impacto en la mejora de toma de decisiones, al suministrar una visión y/o perspectiva mucho más completa y minuciosa de los datos, puede ayudar a los líderes de cada programa a tomar decisiones bien fundamentadas, efectivas y estratégicas en los distintos niveles de la entidad; por otro lado el aumento de la eficiencia operativa puede ser clave para la entidad, ya que con ello los procesos internos que se tengan pueden optimizarse y mejorar la calidad de la prestación de los servicios. Además, es posible fijar un equipo de trabajo que colabore dentro de la plataforma con el objetivo de establecer una mesa de trabajo en el software y cada uno de los miembros pueda aportar algún dato que mejore la comunicación interna de la organización.

Sin embargo, esto puede mejorar la capacidad de investigar y proveer datos con mayor precisión lo que lleva a mejorar la reputación de la entidad y a aumentar la credibilidad de esta. De la misma forma se puede establecer un impacto de carácter tecnológico, ya que se puede presentar una integración de diferentes orígenes de fuentes de datos para poder trabajar en conjunto con mucha más cantidad de información y de este modo obtener un resultado mucho más completo, la

obtención de análisis geoespaciales ayuda a descubrir patrones de tendencia y relaciones ocultas en la información, en adición a ello maneja una automatización en tareas lo cual beneficia al ahorro de tiempo y la eficiencia en los procesos de trabajo; finalmente contiene una interfaz intuitiva lo cual facilita su uso y permite reducir la curva de aprendizaje ante cualquier usuario que interactúe con ella.

6. Objetivos

6.1.General

- Realizar un análisis estadístico avanzado mediante el software ArcGIS Insights, con el fin de descubrir relaciones y correlaciones entre diferentes variables.

6.2.Específicos

- Identificar los algoritmos y técnicas que se deben tener en cuenta para la validación y extracción de datos de diferentes fuentes mediante codificación en JavaScript dentro de los formularios de ArcGIS Survey123.
- Desarrollar una interfaz intuitiva en ArcGIS Insights, creando un modelo que permita agregar o unir datasets del análisis espacial, análisis de datos y aplicación de estilo, con el fin de proporcionar al usuario, una mejor comprensión que facilite el proceso de ejecución del proyecto.
- Generar informes con ayuda de las referencias rápidas, de cada una de las variables que arroja el programa, con el fin de obtener un mejor entendimiento de lo que se quiere presentar.
- Establecer un análisis de tendencias de alto impacto en las diferentes zonas geográficas del país, con el fin de identificar la dirección o cambios a largo plazo en los datos.

7. Marco teórico

Dentro de la investigación que se ha llevado a cabo, se encuentran diferentes fuentes que soportan el objetivo de ArcGIS Insights en la globalidad de plataformas digitales; esta herramienta se basa en tres aspectos principales que son el pilar de funcionamiento de manera integrada y completa: la exploración de datos, el análisis de datos y por último la comunicación que se puede establecer entre los resultados obtenidos. Respecto a la exploración de datos, esta viene siendo la fase inicial en cualquier objeto que requiera de análisis de datos, y por supuesto es la clave para comprender el origen, naturaleza y alcance de la información útil disponible, sabiendo que se puede importar información de diferentes fuentes incluyendo bases de datos, hojas de cálculo, archivos en CSV (Valores separados por comas), KML(Keyhole Markup Language) lenguaje para representar datos geográficos en tres dimensiones, entre algunos otros sistemas de datos espaciales. En cuanto al análisis de datos, se entra a una etapa en donde las técnicas de análisis juegan un papel importante, acá se deben aplicar técnicas de análisis y estadística para obtener resultados como los patrones y relaciones en los datos, el punto a favor con el que cuenta este software es la variedad de herramientas en análisis, desde el área espacial hasta el área de redes lo que conduce a la posible realización de análisis de datos predictivos y el modelado.

Posteriormente se habla de la comunicación que se puede tener entre cada uno de los resultados, este viene siendo una parte esencial para el análisis de datos, para que de acuerdo con ello se compartan diferentes puntos de vista o si es el caso en conocimientos que ayuden a tomar decisiones, creando informes interactivos o tableros de control que contengan gráficos y mapas personalizados, será la manera más efectiva para comunicar lo obtenido visualmente atractivo. Por otro lado, cuando se habla de análisis de patrones y tendencias de alto impacto, hace referencia a

la identificación y estudio de los diferentes patrones y las tendencias más significativas en los datos que puedan tener un alto impacto en una zona o área determinada.

El análisis de patrones implica la identificación de estructuras repetitivas, relaciones o características distintivas en los datos. Estos pueden manifestarse en distintas formas, como agrupaciones, distribuciones, secuencias o correlaciones. Al comprender y analizar estos patrones, se pueden obtener conocimientos útiles sobre el comportamiento de los datos y las relaciones entre las variables (Esri., s.f.).

El análisis de tendencias se enfoca en identificar las direcciones o cambios a largo plazo en los datos a lo largo del tiempo. Esto implica detectar patrones de crecimiento, fluctuaciones cíclicas u otras trayectorias en los datos a medida que evolucionan en el tiempo. Al comprender las tendencias, se pueden tomar decisiones informadas sobre acciones futuras o ajustes en las estrategias existentes (Esri., s.f.).

Dentro de las herramientas que maneja el programa, se encuentra gran variedad de elementos visuales analíticos que se pueden implementar según lo que se quiera observar, por un lado, están los **mapas base**, estos se encargan de proporcionar entornos geográficos de acuerdo con el contenido que se quiere presentar, cuenta con la opción de seleccionar la variedad alojada dentro del programa, incluyendo las opciones que representan datos aéreos, carreteros o topográficos. Por otro lado, los **mapas de bins** resumen datos de puntos, aquellos que provienen de sumatorias, conteos o medias, se usan especialmente para los datos que representan una gran cantidad de puntos o puntos con alta densidad de población.

Además de ello, otra herramienta para analizar los datos son las **zonas de influencia**, aquellas que rodea a una entidad geográfica de un mapa, medida bien sea en unidades de distancia o tiempo.

Otro de los medios por lo que se puede analizar y encontrar buenos resultados son por los llamados **mapas de calor**, donde se identifican la mayor cantidad de puntos en zonas determinadas mediante colores, las áreas con mayor intensidad en color son las que indican la densidad de puntos más alta. Sin embargo, es importante mencionar el método de **normalización**, por el cual es posible convertir los datos totales de recuentos en un índice, radio o proporción que posea una relación con un denominador común.

Los tipos de estadística juegan un papel importante a la hora de presentar la información, dentro de ellos se presenta el **average**, el cual también se conoce como el valor medio o el resultado de dividir la suma de todos los valores que se tengan; el **recuento** como la cantidad de valores que se encuentran en un campo, la **suma, el máximo y el mínimo**.

ArcGIS cuenta con un sin número de técnicas para el análisis de los datos, para comprender la **cantidad** de los datos (gráfico de barras y de columnas, barras apiladas, mapa de árbol, gráfico de burbujas, grafico de calor y tarjeta de indicador clave de rendimiento KPI, símbolos proporcionales, mapas de agregación), para comprender las **relaciones** entre los datos (diagrama de cuerdas, gráfico de dispersión, matriz de dispersión, gráfico de vínculos, mapas de flujo y de telaraña), para comprender el **cambio** de los datos (grafico de barras, gráfico de calor, reloj de datos, gráfico de líneas y serie temporal, combinado, mapa de bins, de densidad) para comprender las **interacciones** entre sus datos (diagrama de cuerdas, gráficos de vínculos, mapa de flujo y de telaraña), para comprender la **distribución** de los datos (Histogramas, diagramas de caja, mapa de ubicación, mapa de valor único, de calor, de densidad, de símbolos, de coropletas, de bins) y para comprender las **proporciones** de sus datos (Gráfico de anillos, mapa de árbol), junto con la creación de gráficos (mapas de agregación, de ubicación, de bins, de valor único, del flujo, entro otros).

8. Estado del arte

En relación con algunas otras indagaciones que se han llevado a cabo, se encontraron diferentes proyectos con distintos enfoques en cuanto a su área de aplicación, entre estos proyectos se encontraron:

El primero se titula “El análisis espacial vincula el aumento de la fiebre del valle con el aumento de la temperatura del mar” dentro de esta investigación se obtuvieron resultados como las aguas más frías de la costa de California, pueden afectar los patrones de viento lo que hace que los vientos más fríos y secos del norte del Pacífico Norte soplan a través del suroeste de los Estados Unidos y secan el suelo. Al determinar la fuerza de las correlaciones entre las tormentas de polvo, la fiebre del valle y los centros de alta población, los científicos también utilizaron modelos estadísticos sofisticados para verificar las conexiones. Validaron sus suposiciones de que tipos específicos de variables climáticas estaban correlacionadas con el aumento de las tormentas de polvo y que las tormentas de polvo estaban relacionadas con casos de fiebre del valle en áreas de alta población (Wright, 2017).

En segundo lugar, se encuentra el proyecto “Mapeo del potencial de energía renovable con la ayuda de NASA POWER”, esta investigación muestra lo que ha logrado la NASA al trabajar con Sistemas De Información Geográfica (SIG), para entregar datos a los usuarios. A través de su proyecto Prediction of Worldwide Energy Resources (POWER), la NASA proporciona mapas interactivos, aplicaciones y servicios de datos que describen las condiciones meteorológicas y la radiación solar y muestran cómo esos recursos pueden estar cambiando con el tiempo, con esto desde que POWER adoptó las capacidades GIS, el equipo ha recibido más de 211 millones de solicitudes de sus datos de 483,000 usuarios únicos, hasta el 28 de febrero de 2021 (Cummings, 2022).

Paralelamente se tiene el proyecto "Los mapas inteligentes guían las investigaciones y acciones de COVID-19 y monitorean la efectividad" este proyecto identifico la capacidad de control de COVID-19 al poder atribuir a las herramientas que cuantifican las entradas de datos y agregan información sobre *quién* y *cuándo* en el contexto de *dónde*. La tecnología que maneja el sistema de información geográfica (SIG) sustenta los paneles de mapeo y proporciona los medios para investigar y comprender la propagación, guiar las medidas de control y evaluar las estrategias para la respuesta a COVID-19, con ello se obtuvieron resultados sobre como los analistas de salud usaron SIG para rastrear la situación a través del espacio y el tiempo. Dando un ejemplo de lo que sucedió en la Universidad Nacional de Seúl, donde un epidemiólogo espaciotemporal llamado Hwang Seung-sik pudo determinar que alguien conocido como "Paciente 31" se convirtió en un super propagador en Corea del Sur. Este paciente infectó a otros miembros de una iglesia, y esto dio lugar al 60 por ciento de los casos actuales del país, a raíz de este resultado es posible crear mapas del nivel de propagación, la velocidad con la que se propaga y las anticipaciones que se pueden llevar a cabo (Este Geraghty, 2020).

Además, se utilizó ArcGIS Insights para crear visualizaciones interactivas que permitieron a los usuarios explorar los datos y los resultados del análisis en detalle, de manera más fácil y entendible hacia un público que posible no tiene experiencia a profundidad en el manejo del software.

Con relación al análisis de data que se ha integrado al sistema de transporte en sus diferentes campos de desarrollo, se tiene Maryland, mapas inteligentes y drones ayudan a los residentes a ver la construcción de carreteras de todo el estado, usando GIS (Sistema de información geográfica) fue posible crear un portal de la Administración de Carreteras del Estado del Departamento de Transporte de Maryland (MDOT SHA), en donde cualquier persona pudiera observar el avance o

supervisar las carreteras con infraestructura en proceso de mejoras, para promover la transparencia y comunicación de los usuarios con datos específicos de ubicación. (Eri blog,2022)

Por otro lado se tiene el análisis del transporte multimodal con base en su crecimiento, esta investigación y proyecto se llevó a cabo en el Foot Traffic de una ubicación específica en U.S. California, el análisis se da para las distintas vías con mayor tráfico de vehículos y demás, para ello crearon un software llamado Predik Data-Driven donde fácilmente se puede utilizar data geoespacial y machine learning, para saber dónde cuándo y porque se produce un accidente, determinar futuros proyectos carreteros y llevar así un registro de patrones de movimiento. (predik-data, s/f)

Dentro de otros modos de investigación y métodos para llevar a cabo diferentes sucesos, se cuenta con la interpolación de datos, en este caso se trata de un modelado de la calidad del agua usando la interpolación, con el fin de analizar niveles de oxígeno disueltos en la bahía de Chesapeake, allí se logró encontrar diferentes niveles de elevación en nutrientes provenientes de escorrentías agrícolas, lo que ligeramente provoca un desequilibrio en los niveles de oxígeno en dicha zona. Para encontrar la interpolación en dicha problemática usaron la herramienta de ArcGIS-Pro, creando y comparando las capas resultantes de los niveles de oxígeno durante el verano de los años 2014 y 2015. (um,s/f)

9. Desarrollo del proyecto

Dentro de esta sesión se establece el enfoque de los métodos y técnicas que se tuvieron en cuenta durante el desarrollo del proyecto, con ello se pudo llegar de manera más fácil a cumplir cada uno de los objetivos empleados, además también facilitó la detección de posibles problemas y limitaciones que surgieron a lo largo del proceso. En este orden de ideas es importante mencionar cada una de las etapas que se manejaron para el desarrollo como tal de la investigación.

9.1. Diagrama de Flujo



Figura 1. Diagrama de Flujo de trabajo

Fuente: Propia

Como todo proyecto es importante saber sobre la metodología de trabajo que se va a presentar, es por ello por lo que la Figura 1, demuestra la secuencia de pasos establecida.

9.2. Planificación y diseño inicial

En primer lugar se llevo a cabo la *etapa de documentación* en donde se realizaron diferentes indagaciones en dirección al ambiente de trabajo que se quería establecer, la manera en que se podía llevar a cabo cada uno de los objetivos y con ello los aspectos que podían repercutir en el proceso de desarrollo a largo plazo cuando ya estuviera funcionando, además se tuvieron en cuenta los antecedentes dentro de la entidad con respecto a investigaciones similares o algunos otros datos que fueran de gran ayuda en la investigación actual. En este caso programas que tuvieran que ver con los siniestros que se presentaban en las carreteras concesionadas de todo el país, y hasta donde la superintendencia pudiera obtener información, este tipo de información suministrada por parte de las concesiones era respaldada mediante imágenes que evidenciaron el suceso y con ello el formato en donde se especificara todo con relación al evento.

De acuerdo con lo anterior se continuo con la etapa en donde se realizó la *identificación y selección de las variables a investigar*, si bien se sabe el proyecto va en caminado a descubrir en donde, como, cuando y porque se denominan sectores críticos de siniestralidad; es por ello que se diseñó un formulario en donde la **ubicación geográfica** fuera clave en el proceso de investigación, y variables como la **gravedad** del evento, la **clase** o tipo de evento que se presentó, las causas del evento y la **condición climática** que se presentaba durante el evento, entre otras variables que se destacan a la hora de tomar decisiones para ejecutar posibles acciones.

Al continuar con el proceso se tuvo que *fijar las fuentes* de las cuales se iban a extraer los datos para poder hacer una comparación, en este caso el origen de la información es la Superintendencia de Transporte, dentro de las bases de datos que diligencia el inspector a la hora de la visita en tiempo real, otra de las fuentes más importantes con la cual se realizaron cruces de datos entre lo que le reportaban a esa entidad y lo que le reportaban a la Superintendencia, es la ANSV (Agencia

nacional de seguridad vial), el INVIAS (instituto nacional de vías), la ANI (agencia nacional de infraestructura) y cada una de las gobernaciones que están a cargo de la vías no concesionadas del país. Estas entidades únicamente comparten información con relación al sitio crítico, el nombre del proyecto y la ubicación del sector en coordenadas, latitud y longitud. La información que se recolecto por parte de estas fue de gran ayuda para direccionar la investigación.

9.3.Creación del Formulario

Dentro de esta sesión se determinó que el método para la recolección de datos seria ArcGIS Survey123, el cual es una herramienta que maneja sistema de información geográfica (GIS), que permite crear y desplegar encuestas bien sea en web o con plantilla de formularios basada en Microsoft Excel, en este orden de ideas se creo un formulario para el programa especial sectores críticos de siniestralidad con plantilla en Microsoft Excel, dentro de el se puede capturar la ubicación geográfica de cada respuesta que diligencie el usuario, además admite varios tipos de medios, como fotos, video y grabaciones de audio según se requiera, proporcionando contexto visual a las respuestas de las encuestas. Survey123 permite aplicar reglas de validación para darle mayor precisión al ingresar un dato, proporcionando campos de carácter obligatorio, con formato especifico, rangos, numero o textos. A continuación, se presentará la plantilla de trabajo en Survey123.

quiere que se presente la información en el *front-end* o en la interfaz después de publicado, si en dos columnas, en lista horizontal o vertical, en desplegable, entre otros.

Dentro de este existe la celda del requerimiento en donde se exige un valor dentro del campo y agregar un mensaje que quiere que se presente en caso de que se salte este espacio, ArcGIS Survey123 nos permite crear un formulario con un amplio numero de herramientas, con la opción de programar en java script y dentro de la encuesta llamar a la función según corresponda, permite poner restricciones en numero de caracteres, en la ubicación, en rangos, entre muchas funciones que hacen que el formulario sea muy completo, el *relevant* se usa para que una pregunta pase a ser visible si cumple con las condiciones de la columna.

Si bien se identifica dentro de esta misma figura 2, en la parte inferior izquierda contiene diferentes pestañas en donde la primera que se titula *Survey*, ahí es donde se diseña como tal el formulario y la que contiene todas las herramientas anteriormente mencionadas, luego esta la de los *Choices* se usaron para especificar opciones de respuesta para preguntas con opción múltiple, para que este funcione se requiere que las columnas de *list_name*, *name* y *label* sean obligatorias. El *external_choices* se usó en este caso para proporcionar valores provenientes de otra fuente, donde las opciones de selección cambian frecuentemente o necesitan estar actualizadas, por ejemplo, allí se almacenaron los nombres de los municipios de todo el país. Por último, el *settings* permite darle un titulo a la encuesta para que de un principio se entienda de que va a tratar, en este caso se le dio el nombre de PESCRI Estadística, el *id* es para darle una identificación única, y con ello el *instance_name* para saber cómo se identifica y se nombran las respuestas individuales.

Luego de que se configura el formulario en el XLSform, se usa el software ArcGIS Survey123 Connect, con el fin de dar una vista previa a como se veria la encuesta en cualquiera de sus presentaciones, es decir, en tablet, movil o pc, tambien se puede brindar un estilo como en el color

de fondo del formulario, el tipo de letra, entre otras. Sin embargo la función principal es previsualizar y publicar, una vez publicado se solicita el diligenciamiento al grupo de enfoque.

PESCRI Estadística

Validación de Vigilado

Código de Proyecto *

NIT *

Información de Supervisado

Código de Proyecto *

Nombre del responsable *

Cargo del responsable *

Correo de Notificaciones *

1 de 1 ✓

Figura 3. Front-end del formulario en Survey123 Connect antes de publicarlo

Fuente: Propia

Como se observa en la figura 3, se presenta el formulario en vista previa a su publicación con el fin de corregir cosas que se quieran cambiar, como el estilo de letra, de organización o la adición de paginas si es el caso, de la misma manera para realizar pruebas en la interfaz con datos verídicos del supervisado para comprobar su funcionamiento.

Por otro lado, cuando se empieza con el diligenciamiento de la información del supervisado, es decir, de la concesión que realiza el reporte, se requiere el código del proyecto, que para cada concesión es único, de la misma manera el número de Nit, una vez diligenciado estos dos datos de manera correcta, el programa realiza una validación con los datos de operadores que se tienen en

el archivo internamente y da paso para continuar con el proceso, en caso de no ser así, se rechaza y no lo deja continuar.

Es importante mencionar que esta validación es llamada como función dentro del formulario, como se muestra en la figura 4.

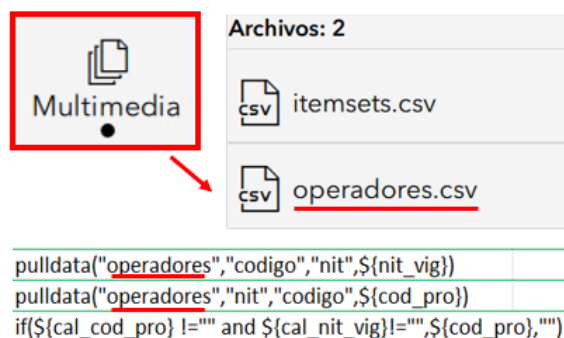


Figura 4. Función usada para extraer los datos del archivo operadores

Fuente: Propia

Además de ello se cuenta con una función que va dada por medio de una java script y es llamada a través de un pulldata en el archivo de XLSForm, como se muestra a continuación.

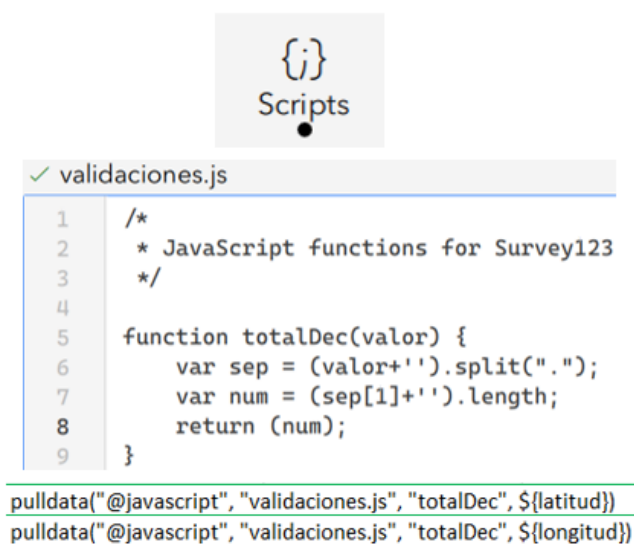


Figura 5. Programación de función en Java Script

Fuente: Propia

De esta manera en la figura 5, es posible observar la función de validación que se creó en javascript dentro del ArcGIS, esta función es para separar un número en dos partes, es decir la parte entera de la parte decimal y con ello la longitud que esta debe tener, esto se usa para validar los valores ingresados en los campos de longitud y latitud, con una extensión mínima de 6 decimales después de la coma, es por ello por lo que dentro del formulario se vuelve a llamar la función.

9.4.Publicación y recopilación de datos

Luego de observar el proceso de creación y todo lo que lleva consigo internamente la encuesta, se procede a mostrar como tal el aplicativo y el software que debe de portar el usuario a quien se le pide por medio de un requerimiento el diligenciamiento del formulario.



Figura 6. Aplicativo ArcGIS Survey123 para el diligenciamiento de la información

Fuente: Propia

La Figura 6, presenta la ventana de inicio del aplicativo que se maneja para el correcto funcionamiento y tratamiento de la información, por un lado si bien se observa es posible iniciar sesión con las conexiones de ArcGIS o con una conexión externa asociada, en este caso la superintendencia cuenta con la licencia de ArcGIS Survey123 por lo que es posible iniciar sesión con cuenta propia de la entidad, al ingresar se presenta una pestaña donde se debe ingresar un usuario y una contraseña, en este caso como el programa especial que se esta tratando es PESCRI deben realizar la entrada con las credenciales correspondientes al programa, cabe aclarar que para cada programa especial que se tiene dentro de la dependencia existe un usuario y una contraseña, en donde todas las concesiones tendrán el mismo pero con diferente código de proyecto. Una vez se ingresa al aplicativo con las credenciales correspondientes nos encontramos con la interfaz, la cual maneja la opción de adquirir la encuesta, dentro del requerimiento se les indica el nombre exacto del formulario que deben diligenciar y tal como se muestra en la figura 7 es como se debe continuar el proceso.

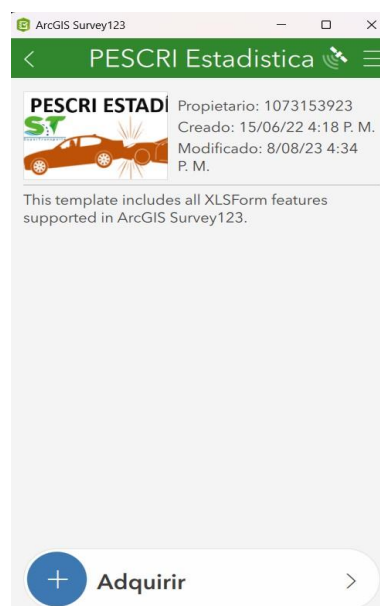


Figura 7. Interfaz ingreso al Formulario

Fuente: Propia

Al oprimir Adquirir, se da un ingreso directo a diligenciar cada uno de los apartados que contiene el formulario y de la misma manera es el indicio de un gran almacén de datos de alta importancia tanto para el país como para la entidad la cual toma decisiones bajo los eventos presentados con material probatorio.

The figure displays two screenshots of the PESCRI Estadística survey form. The left screenshot shows the 'Validación de Vigilado' section with fields for 'Código de Proyecto', 'NIT', 'Código de Proyecto', 'Nombre del responsable', 'Cargo del responsable', and 'Correo de Notificaciones'. The right screenshot shows the 'Validación de Vigilado' section with fields for 'Código de Proyecto', 'Entidad', 'Tipo de servicio', 'Nombre del Proyecto', and 'Nombre / Razón social del Operador'.

Field	Value
Código de Proyecto *	CVN04
NIT *	[Redacted]
Código de Proyecto *	CVN04
Entidad *	ANI
Tipo de servicio *	ICC
Nombre del Proyecto *	Asesoría Técnica Municipal
Nombre / Razón social del Operador *	Asesoría Técnica Municipal

Figura 8. Prueba de encuesta PESCRI

Fuente: Propia

De acuerdo con la figura 8, se tiene lo que viene siendo la verificación del funcionamiento con datos de un proyecto al que se vigila dentro de la entidad, en este punto se obtienen buenos resultados si se considera que las validaciones que se hicieron dentro del programa se tienen en cuenta, ya que como se observa funcionó el patrón de comparación que se programo anteriormente, donde si el Nit y el código eran digitados correctamente se extraía la información del archivo

externo correspondiente al nombre de la entidad, el tipo de servicio que este proporciona, el nombre del proyecto y la razón social del mismo.

The figure displays two versions of a survey form for accident data. The left version is a desktop interface titled 'DATOS DEL SINIESTRO' with a sidebar menu containing sections like 'Información Básica', 'Coordenadas y localización', 'Cabeceras Municipales del siniestro', 'Jurisdicción', 'Localización Siniestro en la Calzada', 'Características de la Vía', 'Gravedad y Clase del siniestro', 'Involucrados', and 'Víctimas'. The right version is a mobile interface titled 'PESCRI Estadística' showing a list of categories: 'Cabeceras Municipales del siniestro', 'Jurisdicción', 'Localización Siniestro en la Calzada', 'Características de la Vía', 'Gravedad y Clase del siniestro', 'Involucrados', 'Víctimas', 'Causas Probables', and 'Ditra / Otra autoridad'. Both interfaces include a map area and a progress indicator at the bottom.

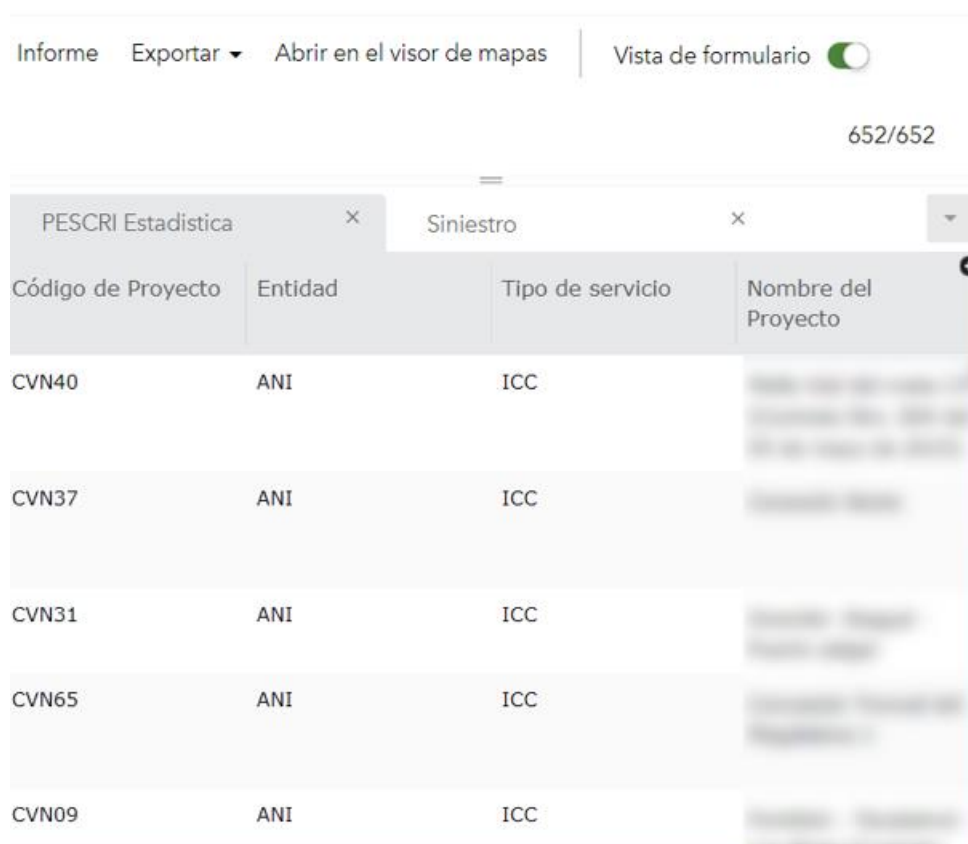
Figura 9. Información sobre el siniestro

Fuente: Propia

Una vez diligenciados los datos del supervisado, se abre una ventana con la cantidad de información que deben reportar mensualmente las concesiones, se encuentran varias sesiones y cada una de ellas con datos de alta importancia como se muestra en la figura 9, los datos del siniestro llevan consigo una particularidad clave y es la ubicación geográfica exacta del evento, es por ello que se exige todo con relación a ello, las cabeceras municipales del siniestro, la jurisdicción a la que pertenece, la localización dentro de la calzada, las características de la vía, la gravedad y clase del siniestro, los involucrados, las víctimas, las causas probables y la información de la dirección de tránsito y transporte (DITRA).

9.5.Almacenamiento y preparación de datos para ArcGIS Insights

Una vez se realiza el proceso de diligenciamiento de la información por parte de cada uno de los vigilados mensualmente, se hace una verificación dentro de la base de datos para presentar un reporte a cada uno de los miembros de la delegatura de concesiones e infraestructura, dentro de los informes que ellos requieren para el reporte, es importante agregar un campo de visualización para la captación de información de manera mucho más elaborada, concreta y fácil de entender, es por ello que se habla de un tablero de control con indicadores que representen la cantidad de datos reportados y todo con relación a ello. Esto es posible observarlo en las figuras 10 y 11.



The screenshot shows a web application interface for the PESCRI database. At the top, there are navigation links: 'Informe', 'Exportar' (with a dropdown arrow), 'Abrir en el visor de mapas', and a toggle switch for 'Vista de formulario' which is currently turned on. A counter '652/652' is displayed on the right. Below the navigation bar, there are two tabs: 'PESCRI Estadística' (active) and 'Sinistro'. The main content area displays a table with the following columns: 'Código de Proyecto', 'Entidad', 'Tipo de servicio', and 'Nombre del Proyecto'. The table contains five rows of data, all with 'ANI' as the entity and 'ICC' as the service type. The project codes are CVN40, CVN37, CVN31, CVN65, and CVN09. The 'Nombre del Proyecto' column is partially visible but mostly obscured by a blurred overlay.

Código de Proyecto	Entidad	Tipo de servicio	Nombre del Proyecto
CVN40	ANI	ICC	
CVN37	ANI	ICC	
CVN31	ANI	ICC	
CVN65	ANI	ICC	
CVN09	ANI	ICC	

Figura 10. Base de datos PESCRI

Fuente: Propia

La figura 10, representa la base de datos de la información diligenciada en el formulario del programa especial de sectores críticos de siniestralidad, dentro de esta tabla es posibles ver lo que ellos contestaron, la fecha de reporte y quien lo hizo, esto con el fin de llevar un control y seguimiento de los eventos en pro de la buena toma de decisiones.

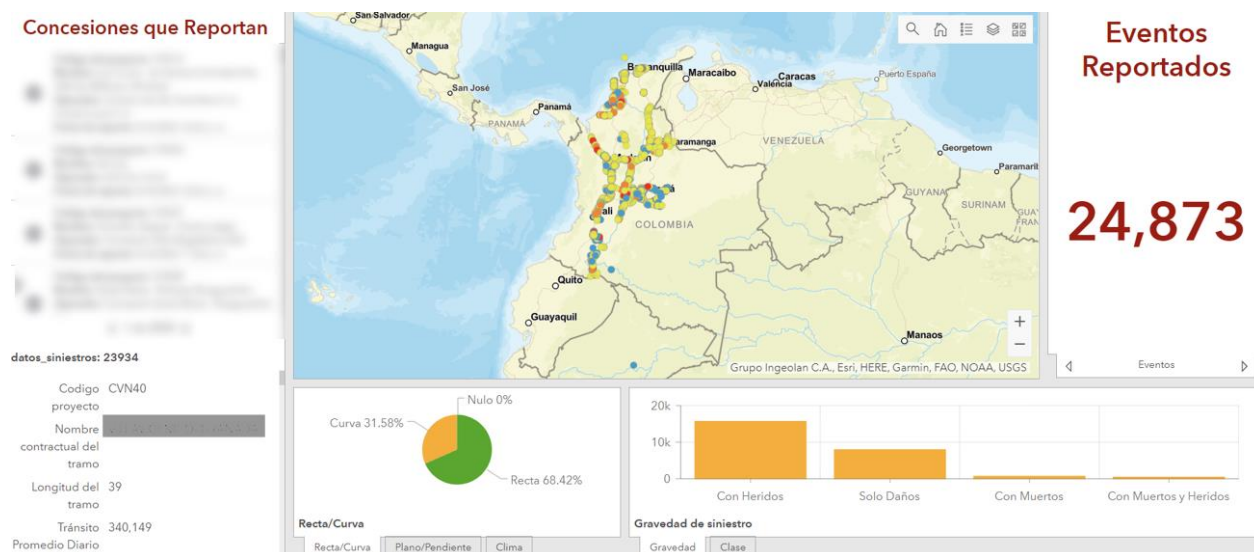


Figura 11. Tablero de control PESCRI

Fuente: Propia

Los datos de una u otra manera se deben presentar por medio de un entorno visual, es por ello por lo que se necesita modelar este tipo de información por medio de un tablero de control donde se refleje la información más importante para toda la investigación, tal como se observa en la figura 11. Además de ello la realización de este tablero es un avance para preparar la información que se va a ingresar a Insights, como el total de los eventos reportados durante el mes, año u otro; la caracterización de la carretera, es decir, si el siniestro ocurrió en curva o en recta y con ello la gravedad del suceso, entro otros datos que ayudan a que la investigación sea muchísimo más completa.

10. Análisis de Resultados y Discusión

Esta sesión se describe bajo aspectos que respaldan y prueban la información recolectada por la superintendencia de transporte al momento de enfrentar alguna situación o evento indeseable, y al tomar una buena decisión, además de ello se presenta un breve comparativo entre otros programas que prestan servicio en analítica de datos como lo son Power Bi y Tableau.

10.1. Comparativo ArcGIS Insights, Power Bi y Tableau

Dentro de este ítem lo que se presenta es una breve comparación entre programas con enfoque de análisis de datos como lo son Power Bi y Tableau frente al software que se usó para el desarrollo del proyecto.

10.1.1. Power Bi Desktop

En primer lugar, se llevó a cabo el desarrollo de un tablero que mostrara los datos de manera concisa usando el programa Power Bi, allí se presentaron datos en las distintas formas de modelado, usando algunas visualizaciones que permitieran un mejor entendimiento de lo que se quería transmitir, dentro de la figura 12, es posible observar la distribución y los datos que se tuvieron en cuenta para el análisis.

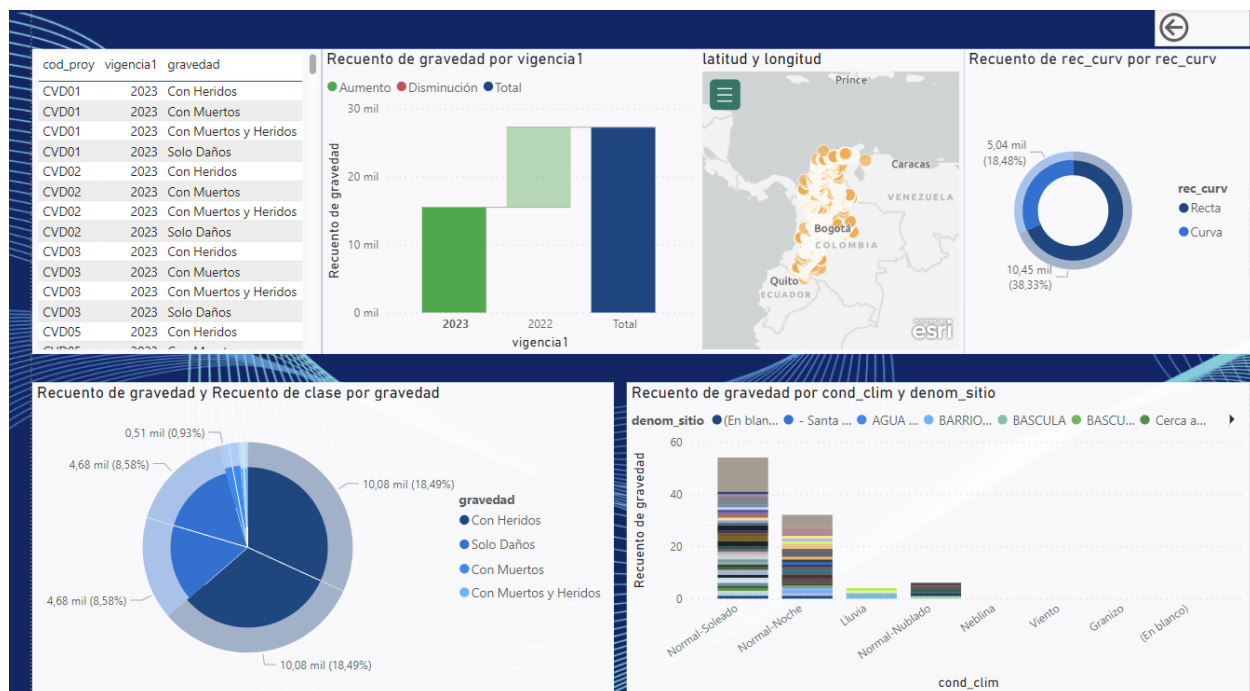


Figura 12. Interfaz gráfica datos PESCRI en Power Bi.

Fuente: Propia

De acuerdo con lo expuesto, es importante mencionar que en primer lugar se muestra una tabla en donde se logra observar el código del proyecto, la vigencia en que este tiene y la gravedad del evento, seguido de ello se cuenta con un gráfico de cascada que permite observar la cantidad de eventos por gravedad con relación al año en que ocurrió que para este caso el enfoque es del 2023. Se tiene un diagrama de anillos en donde se refleja la cantidad de accidentes presentados tanto en recta como en curva, esto dado en milésimas y porcentaje, así mismo se cuenta con un gráfico circular en donde se toma la clase de gravedad y su cantidad. Finalmente se cuenta con un diagrama de columnas apiladas el cual presenta la gravedad por condición climática y la denominación del sitio en donde ocurrieron los hechos.

10.1.2. Tableau Desktop

A continuación, se presenta la prueba que se realizó con los datos de sectores críticos de siniestralidad en Tableau, en donde se logró crear un tablero de control que mostrara cada uno de los datos con mayor importancia para finalmente lograr establecer relaciones y correlaciones entre los datos.

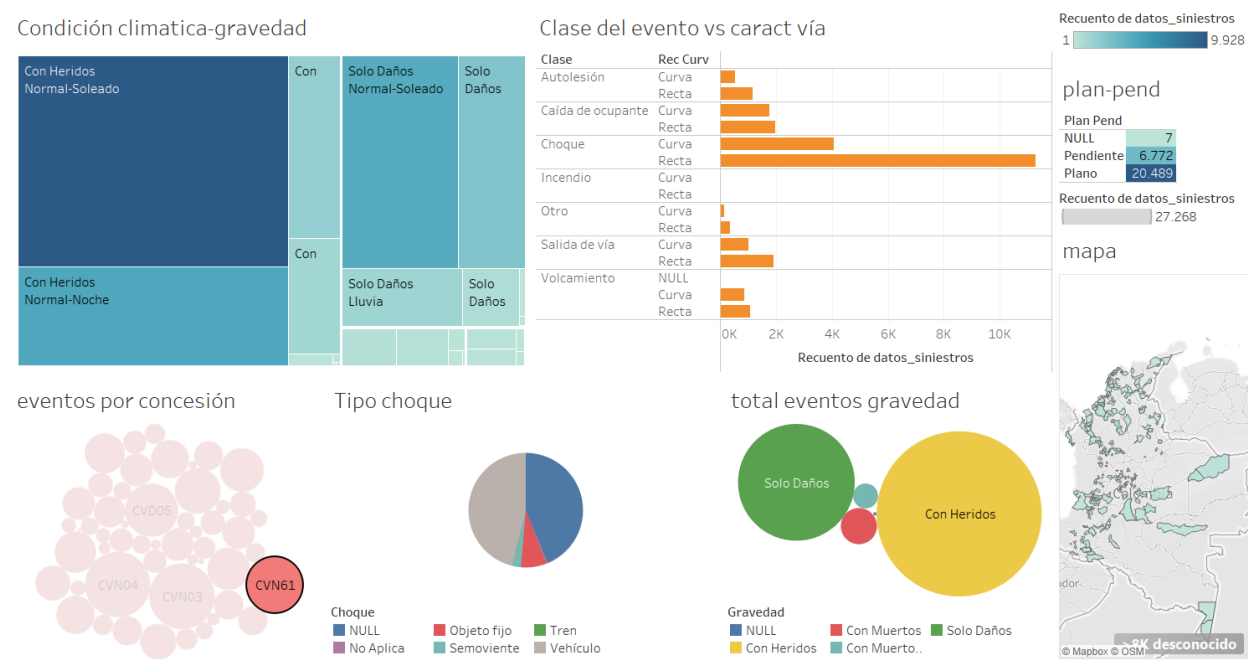


Figura 13. Interfaz gráfica datos PESCR en Tableau.

Fuente:Propia

La figura 13, representa la interfaz gráfica de los datos en el programa Tableau Desktop, en donde se realizaron pruebas con los mismos datos para observar la viabilidad que se podría obtener frente al uso de los demás programas, allí es posible observar los datos de la condición climática en relación al nivel de gravedad, la clase del evento con respecto a la caracterización de la vía, es decir si paso en curva o en recta, en plano o pendiente; los eventos reportados por concesión, el

tipo de choque que se dio y el total de los eventos según la gravedad, sin dejar de lado la ubicación geográfica que el pilar de la investigación.

10.1.3. Fortalezas y debilidades entre ArcGis Insights, Power Bi y Tableau

- En el caso de Tableau tiene la capacidad de integración con variedad de fuentes de datos, puede integrarse con sistemas empresariales por medio de conectores o APIS, mientras que Power Bi se puede integrar con otras herramientas de Microsoft, como Excel, SharePoint y Azure. Esto facilita la integración en un entorno empresarial que ya utiliza productos de Microsoft. ArcGIS Insights se especializa en datos geoespaciales, lo que significa que se integra naturalmente con información geográfica, mapas y capas de ubicación
- Por otro lado, trabajar con Tableau lleva un poco más de proceso, es decir, en primer lugar, para poder visualizar en un tablero de control todo lo que presenta, se debe hacer cada gráfico en una hoja por aparte, este programa no permite en una sola hoja sacar todas las tablas. En este orden de ideas toca estructurar página por página para finalmente poder crear un tablero en donde se presenten todas. Se creó en un principio la tabla de condición climática con relación a la gravedad, y luego en diferentes páginas las demás, para finalmente mostrar la información en el tablero, lo que si se compara con Power Bi y ArcGIS Insights se ahorra todo este proceso.
- Tableau tiende a enfocarse más en la visualización de datos y puede requerir una preparación previa de los datos para obtener el máximo provecho de la herramienta, por otro lado, Power BI ofrece capacidades de modelado de datos más avanzadas, lo que significa que puedes realizar transformaciones de datos directamente en la herramienta.
- Tableau es conocido por su escalabilidad y rendimiento en el manejo de grandes conjuntos de datos y empresas a gran escala y Power BI también puede manejar conjuntos de datos

considerables, pero es posible que se requieran licencias específicas para casos de uso muy grandes y complejos.

- Al usar Power Bi si bien se observa en el tablero de control que corresponde a este programa, es importante verificar que el mapa que está disponible para su uso es suministrado por Esri, la empresa que maneja el software ArcGIS, lo que quiere decir que trabajar con ArGIS sigue siendo mucho más efectivo si el enfoque se maneja bajo un sistema de georreferenciación.
- Las tres herramientas están diseñadas para el análisis y visualización de datos que permiten crear informes y paneles de control de manera intuitiva, la diferencia entre cada uno de ellos es el enfoque que manejan, en el caso de Power Bi es un programa que se usa principalmente en visualización y análisis de datos empresariales en general, mientras que ArcGIS Insights se centra en el análisis y visualización de datos geoespaciales avanzado, permite integrar datos geográficos, crear y personalizar mapas interactivos, permite realizar análisis temporal visualizar su cambio en el tiempo, y realizar geo codificación para convertir direcciones o ubicaciones en coordenadas geográficas.
- De otra manera Tableau tiene un enfoque más general, no se limita a datos geoespaciales, puede conectarse a fuentes de datos sin restricciones geográficas, no necesita de tener conocimientos en programación, ni ser usada por especialistas en el tema. Mientras que Insights se enfoca en el análisis de datos geoespaciales y está diseñado para profesionales que trabajan con datos geográficos.
- Usar ArcGIS Insights para llevar a cabo la analítica de datos establecida, ha sido una de las mejores decisiones que ha tomado la superintendencia de Transporte, teniendo en cuenta que el enfoque que se tiene es trabajar con sistema de información geográfica ArcGIS

Insights cumple las expectativas y los objetivos de la misma, a futuro se pueden desarrollar propuestas ligadas a la inteligencia artificial para obtener de formas más concisas los sistemas de información en tiempo real y bajo material probatorio como cámaras que capten videos de los hechos ocurridos y bajo qué condiciones, ayudando a tomar decisiones de manera mucho más informada y segura.

10.2. Análisis de datos ArcGIS Insights

Al preparar los datos recolectados para ingresarlos a la etapa de análisis es importante saber con claridad que es lo que se quiere visualizar, presentar y transmitir. Como primer paso, dentro del libro creado para el programa especial de sectores críticos de siniestralidad, se presenta la interfaz gráfica que se realizó, figura 12.



Figura 14. Interfaz gráfica PESCRI en ArcGIS Insights

Fuente: Propia

Dentro de la interfaz gráfica que se desarrolló en ArcGIS Insights, en la parte inferior izquierda, se cuenta con diferentes ventanas que se categorizan como datos por Fecha, por Gravedad, por características de la carretera y por patrones y tendencias, tal como lo presenta la figura 14.

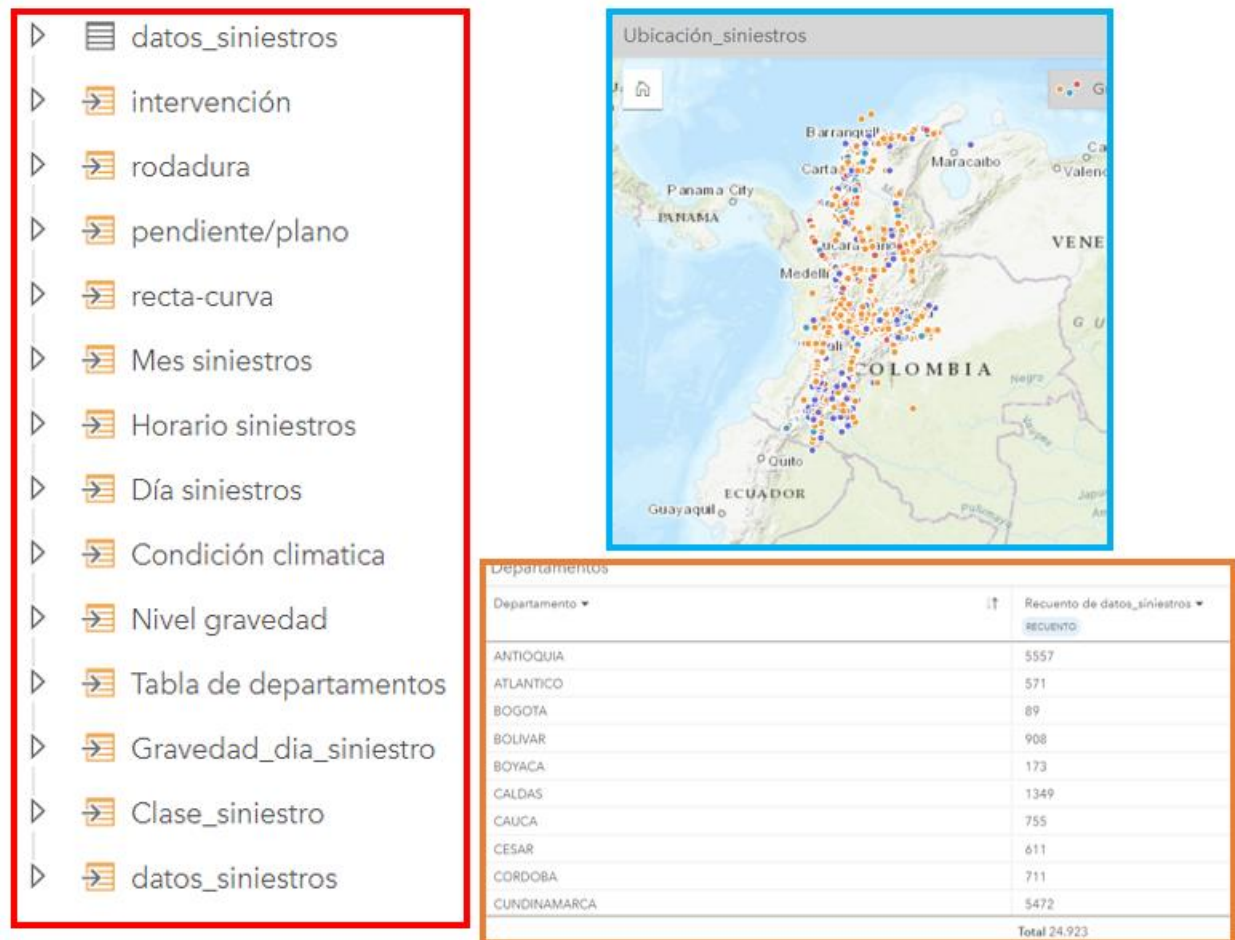


Figura 15.Ventana Mapas

Fuente: Propia

La ventana titulada como mapas figura 15, muestra todos los datos en general, lo que encierra el recuadro de color rojo, son los datasets, es decir, el campo donde se almacena un conjunto de datos, en este caso datos de carácter numérico, textual, auditivo y/o visual. Se presentan datos como la clase o tipo de siniestro que ocurrió, la gravedad del evento según el día, el departamento en donde paso, y variables que hacen de la investigación algo mucho más completo.

El recuadro azul que encierra el mapa destaca la ubicación geográfica de los eventos ocurridos dentro del país, donde se puede filtrar por año, y cada punto maneja una ponderación por color, como se observa en la figura 16.

Gravedad	CALCULAR
• Con Heridos	15.796
• Con Muertos	760
• Con Muertos y Heridos	360
• Solo Daños	7978

Figura 16. Ponderación por color

Fuente: Propia

Por último, el recuadro de color naranja que rodea la tabla de datos destaca la cantidad de siniestros que se presentan por departamento y la sumatoria de todos para encontrar un resultado final.

10.2.1. Datos por fecha

De la misma manera a lo anterior, se exhibe la ventana en donde se almacenan los datos por fecha, acá se estableció un análisis de forma que se encontraran relaciones y correlaciones entre cada uno de los datos recogidos, como se revela a continuación.

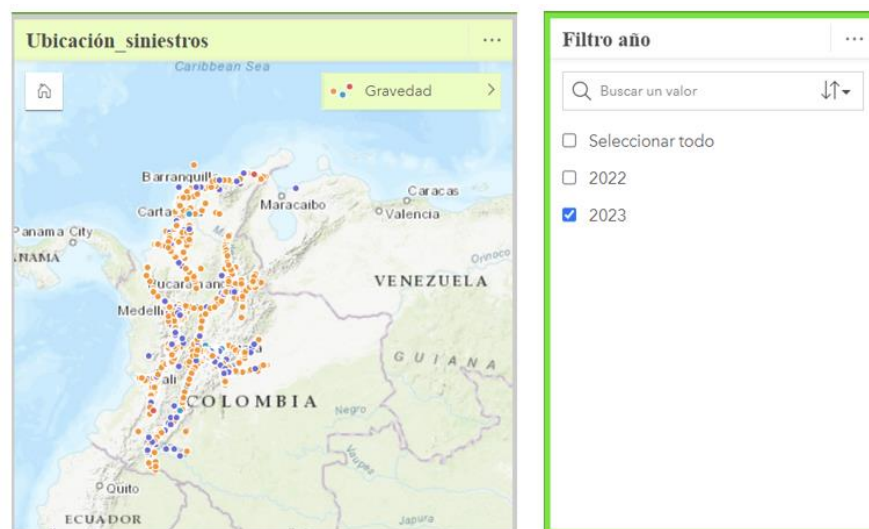


Figura 17. Mapa de ubicación de los siniestros ocurridos en el año 2023

Fuente: Propia

Simultáneamente lo primero fue realizar un filtro de los años para que en este orden de ideas se facilitara obtener información específica, al poner la opción de 2023 inmediatamente el cambio se hace en el mapa, allí se muestran los eventos que se han presentado en las diferentes regiones del país, con coordenadas de ubicación exactas y con el nivel de gravedad según el acontecimiento.

Una vez conocidos estos datos se procedió a relacionar diferentes variables que explicarían más a detalle lo que sucedió y las causas de esto, como el mes, día exacto, la hora y la gravedad.

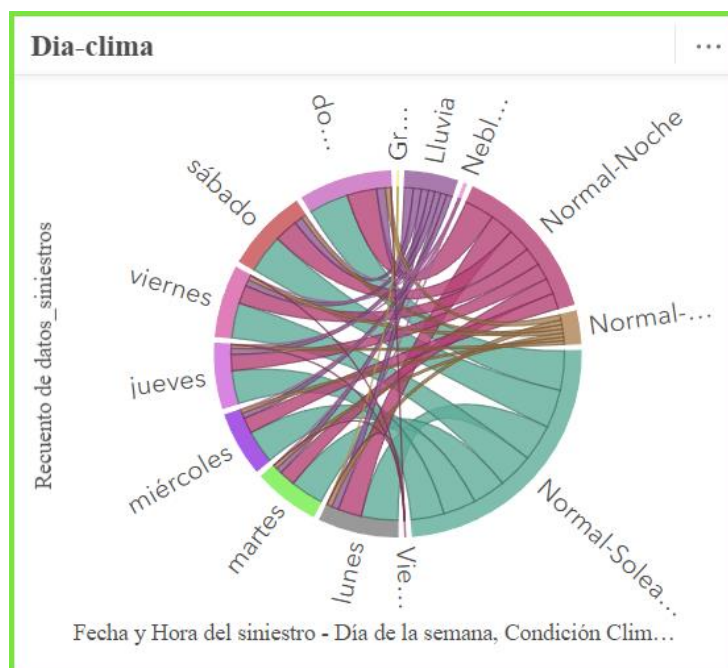


Figura 18. Diagrama de cuerdas día de la semana respecto a la condición climática

Fuente: Propia

La información que se presenta por medio del diagrama de cuerdas, realiza una comparación del día de la semana respecto a la condición climática que se presentaba a la hora del evento, de acuerdo a la figura 18, se obtiene que en un día normal soleado es cuando más se presentan accidentes, con una cantidad de 6396 en cualquier día de la semana, y en los que menos se presenta es cuando cae granizo, neblina y viento, por otro lado en condición normal noche, también existe un numero considerable con 3697 hechos, al igual que normal nublado con 759 y con lluvia 1233. Ahora bien, se es posible irnos más a detalle con los días específicos en donde ocurren acontecimientos bajo condiciones climáticas específicas, por ejemplo, con viento 8 eventos y solo hay cinco días en donde se han registrado las anomalías, en este caso domingo, viernes, jueves, martes y lunes; lo que dicta que ni los sábados ni los miércoles se ha presentado algún inconveniente bajo viento.

Bajo condiciones de caída de granizo solo se han reportado 2 en lo que va del año, con días domingo y lunes y bajo condición de neblina se reportaron 56 en todos los días de la semana.

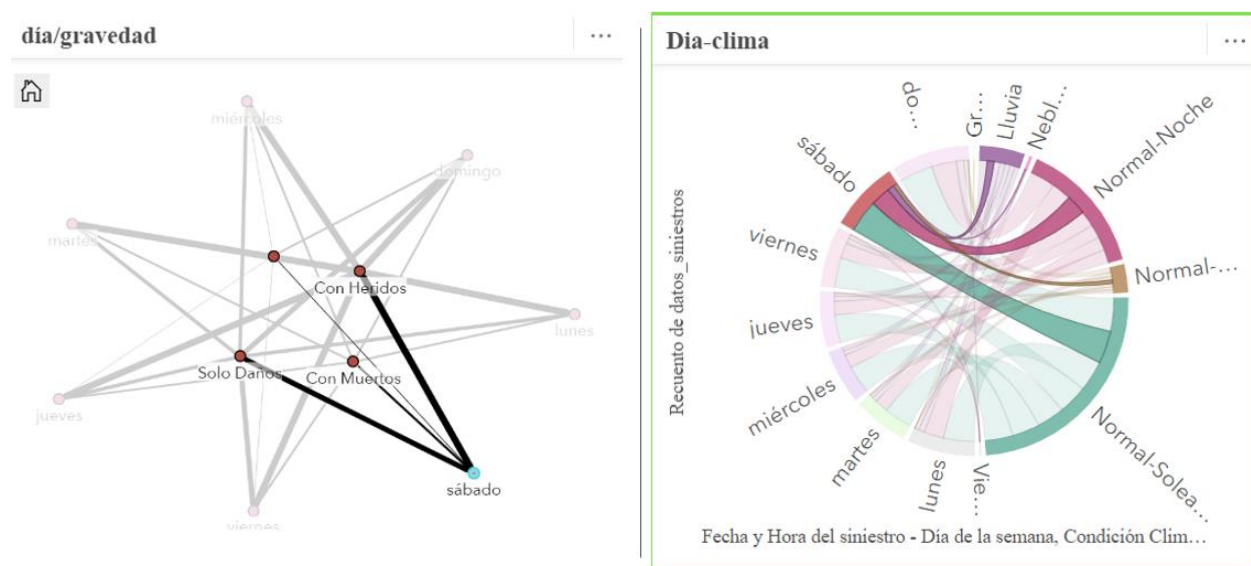


Figura 19. Relación y correlación entre Día/gravedad y Día/condición climática

Fuente: Propia

Encontrar relaciones y correlaciones entre los datos es algo que pone al análisis en su punto más alto, es tan esencial contar con una herramienta que nos permita ver de que manera repercute el uno con el otro, siendo así, la relación que se observa en la figura 19, da a conocer por un lado los datos del día versus la gravedad del mismo mediante un diagrama de vínculos y por otro lado se muestra el día de la semana respecto a la condición climática que se estaba presentando en el momento expuesto a través de un diagrama de cuerdas.

Entre los resultados que se pueden obtener a partir del gráfico, es la correlación que existe entre ellos, si bien se percibe al seleccionar los sucesos que pasan el día sábado, este inmediatamente expone los datos con relación al mismo, es decir, enseña la gravedad del siniestro en el diagrama de vínculos, manifestando que lo más común es que dentro del evento la gravedad de este sea solo con daños o con heridos y que con muertos y heridos los eventos disminuyen, mientras que en el

diagrama de cuerdas, el día sábado la mayor cantidad de accidentes se presentan bajo condiciones normal soleado y normal noche, otros pocos bajo lluvia, neblina y normal nublado.

Al entrar a detallar diagramas como estos es posible que se obtengan un sin numero de resultados, y todos de alta importancia, en este caso datos con relación a la fecha y hora del acontecimiento son más.



Figura 20. Mapa de calor horario y día del siniestro

Fuente: Propia

Acorde con el mapa de calor expuesto en la figura 20, existe una relación entre sus datos al momento de querer saber que evento se presentó durante el jueves y la gravedad del mismo, automáticamente en el otro diagrama se resaltan los horarios en el momento que ocurrió el incidente bajo la misma condición, con muertos y heridos, de acuerdo a ello estos hechos se suelen frecuentar en horarios entre las 00:00 am y 05:00 am y nuevamente de 08:00 am a 12:00 pm, de 2:00 pm hasta las 6:00 pm y de 9:00 pm a 11:00 pm, lo que en teoría es a toda hora del día y la noche a excepción de horas en las que la gente desayuna, almuerza y come.

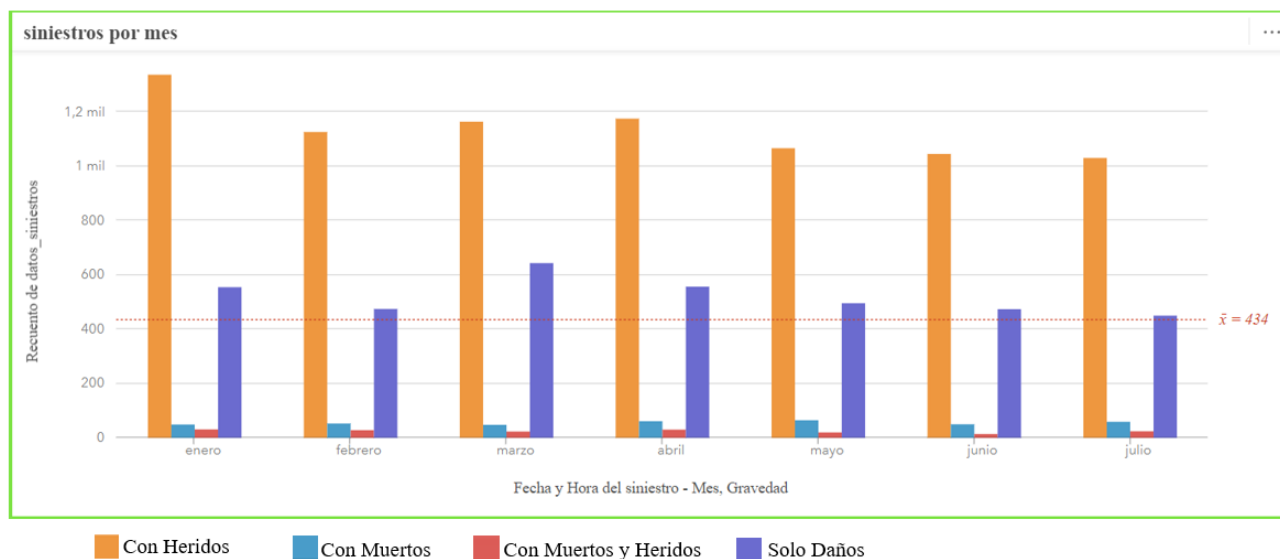


Figura 21. Diagrama de barras siniestros por mes

Fuente: Propia

En cuanto a los eventos que se han presentado en lo que se lleva del año hasta el mes de julio, los datos que se obtuvieron vienen dados por la gravedad de este, es decir la figura 21, dice que en enero del presente año la cantidad de accidentes con solo heridos sobrepasa los 1200, con muertos es aproximadamente $\frac{1}{4}$ de 200, con muertos y heridos es menor de los 40 hechos aproximado, y en solo daños sobrepasa los 500.

Además de lo anterior, es evidente que durante los meses siguientes como febrero, marzo, abril, mayo, junio y julio los datos son semejantes, lo que como resultado la curva que se establece por eventos bajo gravedad con solo heridos es la mas alta, oscila entre los 1000 y 1300; al igual que los demás, siendo solo daños el segundo lugar entre los 450 y 650, bajando hasta los hechos con muertos entre los 45 y 65, dejando por último a aquellos con muertos y heridos entre los 15 y 30.

A su vez, este grafico nos permite identificar cual ha sido el mes con más demanda en sucesos según la categoría, enero con el mes más alto en numero de heridos y julio con el mas bajo, marzo con el más alto numero en daños y julio con el menos, abril y mayo con el más alto en numero de

muertos y marzo junto a enero con el menor número de fallecidos, así mismo el mes con más cantidad de hechos con los dos, muertos y heridos, enero, febrero y abril y el que menos tiene viene siendo junio.



Figura 22. Análisis Relación y correlación de datos por fecha

Fuente: Propia

Para finalizar con el análisis de los datos por fecha, se presenta la figura 22, dentro de ella se encuentra un análisis de relación y correlación, al seleccionar un dato de la tabla titulada siniestro por mes, todos los datos relacionados a este, inmediatamente cambian en las demás, en este caso se realizó la selección para aquellos eventos en donde hubieron fallecidos en el mes de enero, dado esto se obtiene que los incidentes se presentaron bajo condiciones climáticas diferentes y en los distintos días de la semana, aunque fueron pocos la gravedad es alta y su ubicación geográfica arroja puntos tanto en la parte central del país como en la parte norte del mismo, como se puede apreciar en el mapa correspondiente a los puntos de color azul.

Esta clase de análisis se hace interesante cuando se maneja cierta cantidad de datos con el fin de observar que lo que afecta a un lado puede afectar a muchos más implícitamente.

10.2.2. Datos por nivel de gravedad

Este apartado describe todo lo que se obtuvo en relación con el nivel de gravedad del acontecimiento presentado, el total de los incidentes y las sesiones según su categoría de riesgo, las relaciones y correlaciones que se encontraron al cruzar los datos con otros y demás.

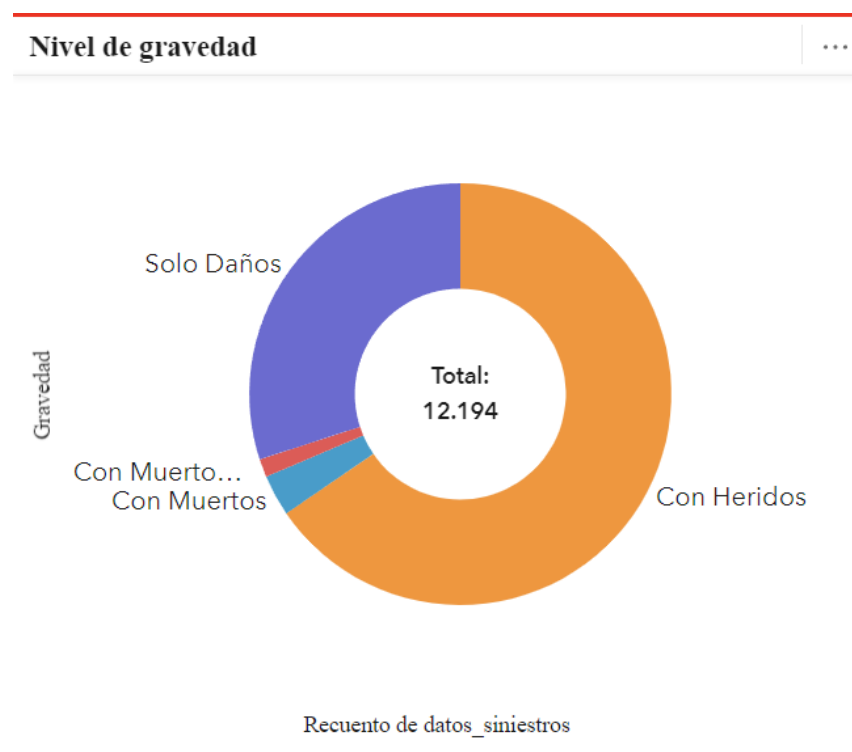


Figura 23. Gráfico de anillos-Nivel de gravedad

Fuente: Propia

Representar el nivel de gravedad del hecho mediante un gráfico de anillos figura 23, es muy práctico porque de esta manera se puede captar con facilidad lo que este quiere transmitir, por un lado, el total de eventos que han ocurridos desde principio de año del 2023, hasta los primeros 10

días de agosto del mismo año, con un valor de 12.194, de los cuales gran proporción son con heridos y daños, y muy pocos con muertos y con muertos y heridos.

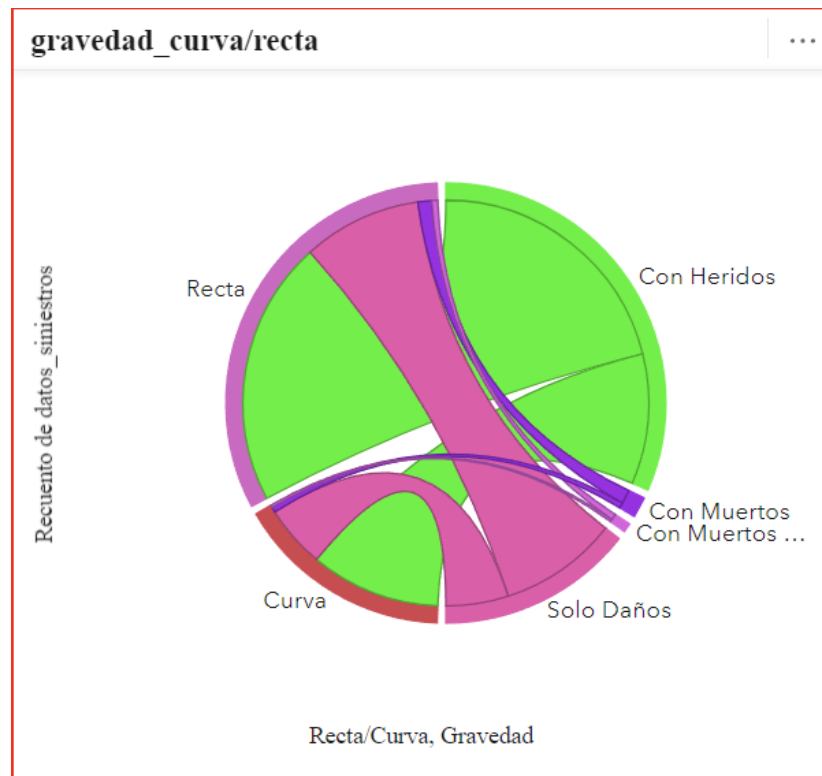


Figura 24. Diagrama de cuerdas- gravedad vs curva o recta.

Fuente: Propia

En lo que sigue se elaboró el análisis del nivel de gravedad del acontecimiento con respecto al diseño geométrico de la carretera, en otras palabras, si este se presentó en curva o en línea recta, según la figura 24, una parte considerable es la que refleja los hechos en recta y con heridos; de otra forma en curva y con heridos también se presentan varios reportes, aunque si se entra a comparar los datos de solo daños en curva y recta también existe una diferencia en cantidad. Ahora bien, si se analizan los datos de muertos en recta respecto a los mismos, pero en curva se presenta con mayor cantidad en recta lo que quiere decir que tal vez al ser recto se puede subir mucho más el nivel de velocidad que lo establecido normativamente, en muchos autos la velocidad no se siente

sino hasta que tiene que frenar y cuando es casi imposible es cuando pasan anomalías, o por la prudencia del conductor, entretenimientos y demás. Estos datos son semejantes a los datos combinados o con muertos y heridos para ser mas claros, estos se dan mayormente en línea recta que, en curva, la gente se relaja un poco más y no piensan en lo que puede llegar a pasar.

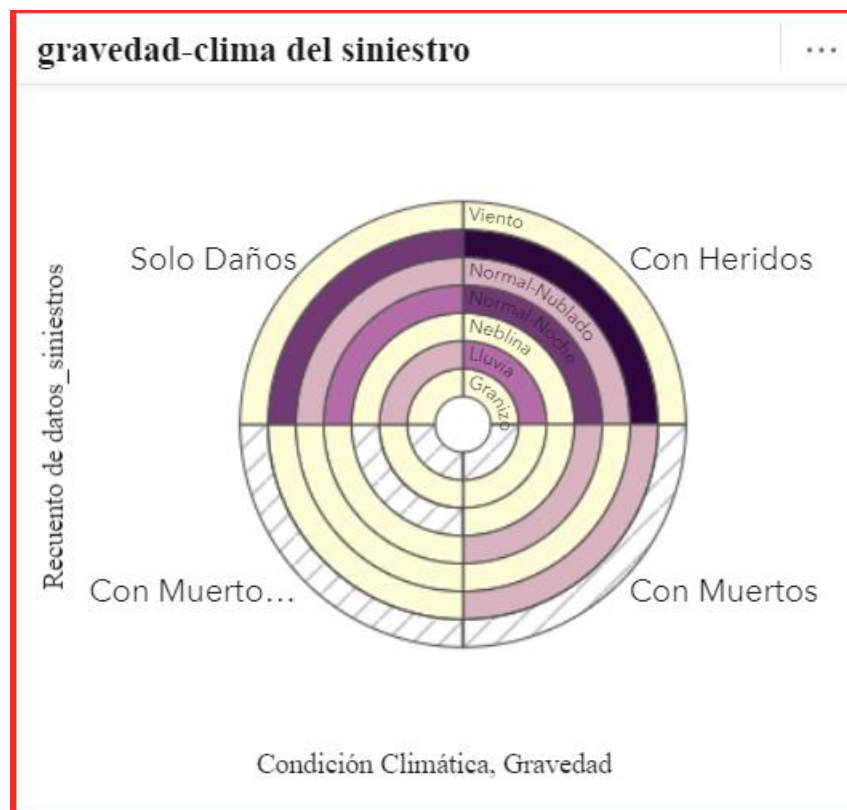


Figura 25. Dato de reloj-Gravedad vs condición climática

Fuente: Propia

Observar los datos de gravedad vs condición climática por medio de datos de reloj figura 25, exhibe que no todos los eventos se presentan bajo las diferentes condiciones climáticas, es decir, bajo viento o con granizo nunca se ha presentado uno con fallecidos, y bajo condición como el viento, la neblina y el granizo tampoco se han presentado hechos con muertos y heridos. De otro modo con solo heridos y con solo daños si se han presentado bajo todas las condiciones climáticas.

Este grafico es fácil de analizar y entender, puesto que a la hora de tener datos como estos la mejor representación viene dada por él.



Figura 26. Mapa de árbol-Tipo de siniestro

Fuente: Propia

De acuerdo con la figura 26, es posible observar un mapa de árbol el cual muestra el tipo o clase de siniestro presentado, de acuerdo con ello es posible evidenciar que la mayoría de los eventos son por motivo de choque, en cifras exactas contiene 6808 choques, como siguiente evento con más cantidad es la caída de ocupantes con 1750, luego va la salida de vía con 1246, volcamiento con 873, autolesión con 856, atropello con 453, incendio con 23 y otro con 185. Este mapa arroja la cantidad exacta de los datos que la componen, desde la mayor hasta el menor, es de fácil interpretación y análisis.

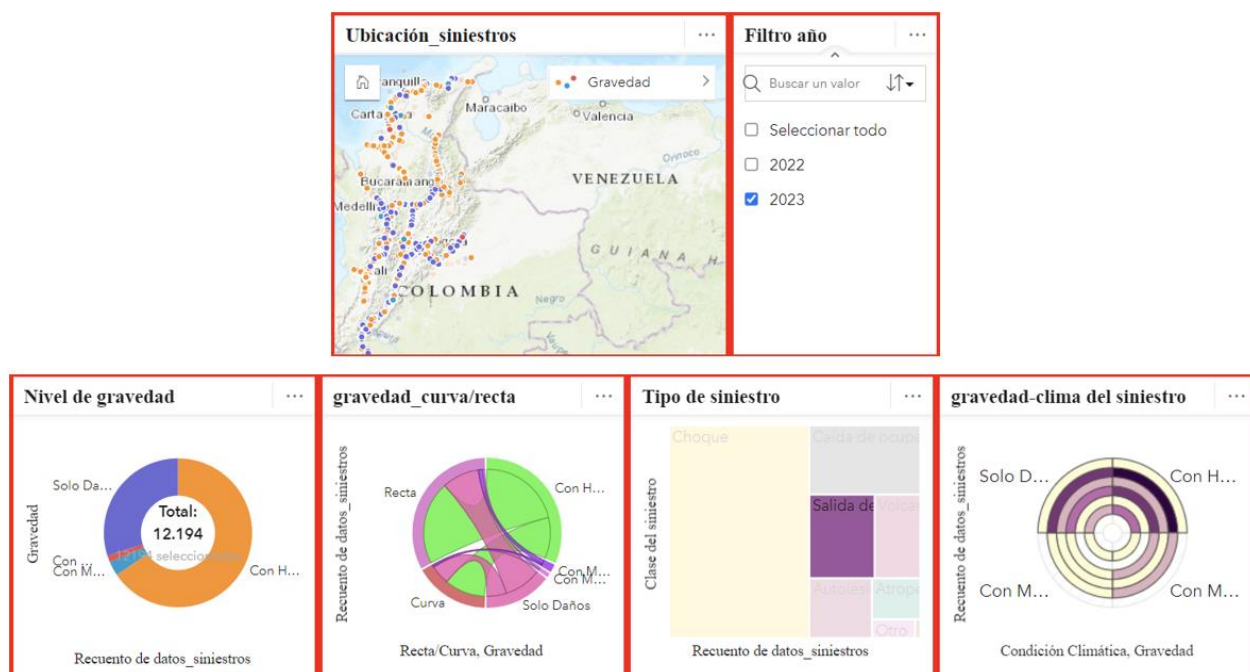


Figura 27. Relación y correlación de los datos por nivel de gravedad

Fuente: Propia

Como se ha mencionado anteriormente, todos los datos se encuentran relacionados entre sí, al querer saber sobre los hechos que presentan salida de vía, se obtiene que se han presentado a lo largo del país en grandes cantidades, y con todos los niveles de riesgo, en curva y en recta y bajo condiciones climáticas de casi todas sus categorías a excepción del viento y granizo. Todos estos datos para el año en curso 2023.

En este análisis se lograron interpretar cada una de los diagramas y gráficos que se presentaron, cada uno de ellos con un objetivo en la manera de exhibir la información, con características propias que pueden llegar a funcionar en muchos más campos de acción, con diferentes enfoques y por supuesto diferentes datos.

10.2.3. Tendencia en los datos

Estos datos presentan valores de carácter repetitivo, aquellos que se pueden observar cómo varían en el tiempo, que tanto varían y hacia donde tienden.

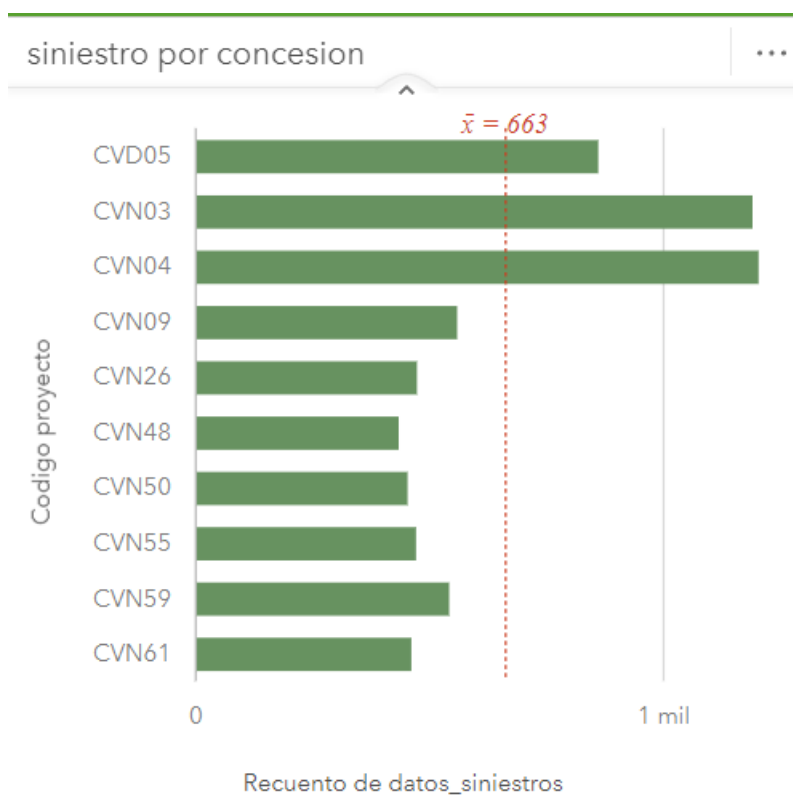


Figura 28. Gráfico de barras, ponderación de siniestros por concesión

Fuente: Propia

La figura 28, presenta los datos de siniestros por concesión, en este caso la CVN04 es la que más siniestros presenta con una cantidad superior a los mil eventos, lo que parece ser que dentro de este tramo hay existencia de trayectos peligrosos que se deben atender con mucha precaución, puede ser por curvas peligrosas, por fallas geológicas o por tramos rectos que demandan el uso de alta velocidad.

Es importante destacar que la CVN04 es el tramo que comunica Armenia- Pereira- Manizales, de acuerdo con ello y con fuentes externas, se dice que es el tercer tramo más mortífero del país y la mayoría de los siniestros en gran cantidad son motociclistas, por motivos de fallas mecánica, llamada por celular entre otras distracciones. En segundo lugar, se encuentra CVN03, lo cual pertenece a la concesión vial de Rio Negro y Medellín y también presenta más de mil eventos, lo que resulta ser semejante a la concesión tratada con anterioridad, donde el nivel y ponderación de la gravedad es alta y donde muy probablemente las razones de accidentalidad son similares, causas por distracción, imprudencia, fallas mecánicas y demás.

Otra de las vías que presentan riesgo va dirigido por la concesión CVD05, la cual representa el tramo de Chía- Mosquera- Girardot y el ramal de Soacha, acá como en las anteriores los datos de siniestralidad son bastante considerables, dado que no pasan de los mil, pero sí de los 800 en lo que resulta algo de atención inmediata, acción para mitigar o evitar que este número, crezca con el pasar del tiempo. Para los demás datos la media se encuentra en 500, lo que quiere decir que al igual que los demás datos presenta un déficit en cosas que requieren atención, puede ser cuestión de señalización, de concentración, de sensores de velocidad, entre otros.

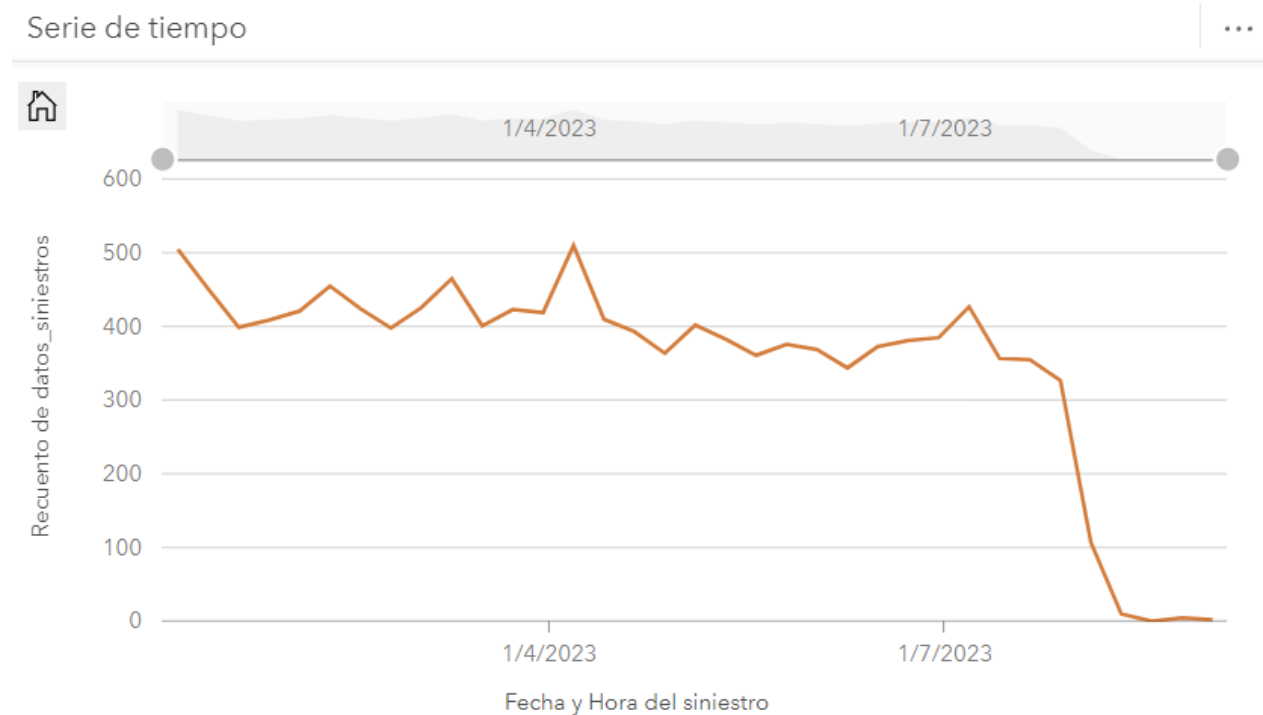


Figura 29. Serie de tiempo siniestros durante el 2023

Fuente: Propia

La representación de los datos por medio de un gráfico de serie de tiempo, es una de las mejores maneras que se tienen para observar el cambio que se va presentando a lo largo del tiempo, análogamente este exhibe la cantidad de siniestros con respecto a la fecha, figura 29, empezando por el 1 de enero del 2023, hasta el 13 de agosto del 2023; en este orden de ideas al identificar el comportamiento no es tan estable, en ocasiones tiende a bajar o tiende a subir, sin embargo el pico más alto sobrepasa los 500 eventos en una fecha que particularmente coincide con semana santa, de tal modo es un momento en el que el flujo del transporte se dispara severamente, se presenta alta demanda de personas que se trasladan de un lado a otro, esto sin tener en cuenta que algunos de los conductores viajan por horas muy largas y las causan que pueden provocar algo no deseado puede ser micro sueños, o agotamiento visual.

Otro de los picos que se pueden apreciar se presenta en inicios de meses como en febrero y marzo; así mismo el ultimo pico que presenta cambio significativo se encuentra en el rango del 2 al 9 de julio, semana en donde posiblemente el momento es vacacional, bien sea de jardines, colegios, universidades, empresas, entre otros, por lo que también la operación de transporte es de alto impacto.

Ahora bien, a lo largo del año los eventos oscilan entre los 330 y los 510, y en este último mes de agosto los que se han reportado han sido pocos, es importante destacar que la fecha de reporte por parte de las concesiones es mensual y se realizan los primero 15 días de cada mes, en este orden de ideas hasta ahora el pico más bajo se encuentra en el mes de agosto con aproximadamente 150 siniestros.

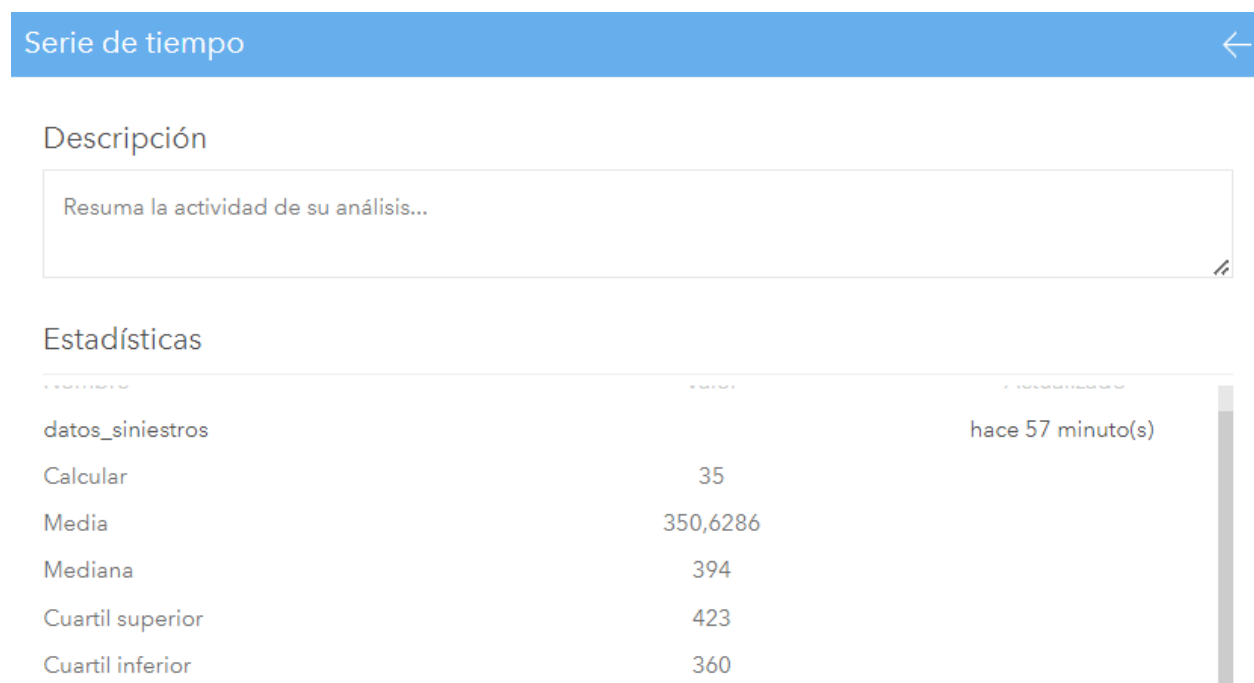


Figura 30. Descripción y estadística de los datos

Fuente: Propia

Dentro de estas tarjeta por la otra cara, es posible realizar una pequeña descripción de lo que se esta presentando, y con ello los datos estadísticos con mayor relevancia que pueden ser clave para comparar datos y tomar acciones o decisiones, en la figura 30, se tiene que la media, la mediana, el cuantil superior y el cuantil inferior; cuando se habla de la media que en este caso es de 350.62 se refiere a la posición que se tiene frente a todos los datos, es decir, el promedio de ellos, la tendencia central del conjunto de datos que se tienen. Esto resulta útil para resumir los datos, para compararlos y comprender como se sitúa en relación con las demás medias, puede ayudar a identificar patrones y tendencias en datos a lo largo del tiempo o en función de otras variables, a menudo se utiliza como base para estimaciones y predicciones en análisis y modelos estadísticos.

Por otro lado, se tiene la mediana con un valor de 394, corresponde al punto medio dentro de la distribución de los datos, esta mediana no se encuentra tan lejos de la media, lo que quiere decir que no existen valores atípicos o extremos que puedan afectar de manera significativa el resultado.

Cuando se habla de cuantil superior es cuando se dividen los datos en dos partes, una que contiene un porcentaje específico para los valores más altos y otro porcentaje restante para los valores más bajos, en este caso el cuantil superior es de 423 lo que traduce a que contiene el 25% superior de los valores y que la otra parte pertenece al 75% inferior de los valores.

Ahora bien, el cuantil inferior es de 360, lo que quiere decir que al menos el 25% de los valores se encuentran en 360 o por debajo de este, lo que indica que una parte de los datos es relativamente baja.

Existen más métodos de sacar estadísticamente los valores, como la desviación estándar, el skewness y el exceso de curtosis, los cuales arrojan datos interesantes que pueden ayudar a observar la tendencia de los datos.

10.2.4. Informes de presentación

El diseño de informes en plataformas que contienen muchísimas herramientas, proveen creaciones de manera automática siempre y cuando se tengan los nombres de las variables establecidos, ArcGIS Survey contiene una herramienta que permite por medio de las referencias rápidas, generar el informe con un diseño de estructura personalizada, o con opción de usar algún modelo preestablecido, en este orden de ideas, se creo una plantilla en Word que llevara a detalle las especificaciones de las actas que se manejan en la entidad, como se ve a continuación.

\$<\$summary>

PESCRI Estadística

pescr_i_estadistica

Observaciones de la inspección

Causas probables	Característica de la vía	Gravedad del siniestro	Número de involucrados
#{datos_siniestros}	{cond_clim}	{gravedad}	{nota_inv}
{causa_infra}	{sup_rod}	{pond_gravedad}	{carga}
{causa_oper}	{rec_curv}	{clase}	{automoviles}
{causa_contri}	{plan_pend}	{choque}	{motocicletas} {/}

Figura 31. Referencias rápidas para el informe

Fuente: Propia

La manera de crear los informes dentro de esta plataforma es por medio de referencias rápidas, figura 31, estas referencias el programa las usa de acuerdo con el nombre de cada variable y como esta se encuentra almacenada dentro del XLSForm, para que estas referencias funcionen se debe poner la etiqueta de inicio `#{datos_siniestros}` y la etiqueta de fin `{/}`, los demás datos van con

el signo de pesos y corchetes, así es como se busca la información almacenada en cada variable y de esta manera se extrae hasta el formato.

PESCRI Estadística

pescr_estadistica

Observaciones de la inspección

<u>Causas probables</u>	<u>Característica de la vía</u>	<u>Gravedad del siniestro</u>	<u>Número de involucrados</u>
Ninguna	Normal-Soleado	Con Heridos	
Exceso de velocidad	Asfalto	0.33	0
Ninguna	Recta	Choque	0
	Plano	Vehículo (Bicicleta, Automóvil, Moto)	1
Ninguna	Normal-Soleado	Con Heridos	
Exceso de velocidad	Asfalto	0.33	0
Ninguna	Recta	Choque	0
	Plano	Vehículo (Bicicleta, Automóvil, Moto)	2
Ninguna	Normal-Soleado	Con Heridos	
Exceso de velocidad	Asfalto	0.33	1
Ninguna	Recta	Volcamiento	1
	Plano		0
Ninguna	Normal-Soleado	Con Heridos	
Exceso de velocidad	Asfalto	0.33	0
Ninguna	Curva	Autolesión	0
	Plano		1
Ninguna	Normal-Soleado	Con Heridos	
Exceso de velocidad	Asfalto	0.33	0
Ninguna	Recta	Choque	1
	Plano		1

Figura 32. Informe reporte de PESCRI estadística

Fuente: Propia

Como se observa en la figura 32, es la manera como queda el informe, si hay existencia de causas probables, aparecerá sino queda como ninguna, las características de la vía indican en que

condición climática se encontraba, si era asfalto o destapado, si era recto y curva, la gravedad del siniestro lo que indica el nivel, y la ponderación de esta, y por último el número de involucrados.

Por otro lado, se puede visualizar un trazo de todos los mapas de análisis que se ingresaron en ArcGIS Insights, las tablas y demás, en caso de que se quiera ver todo el orden.

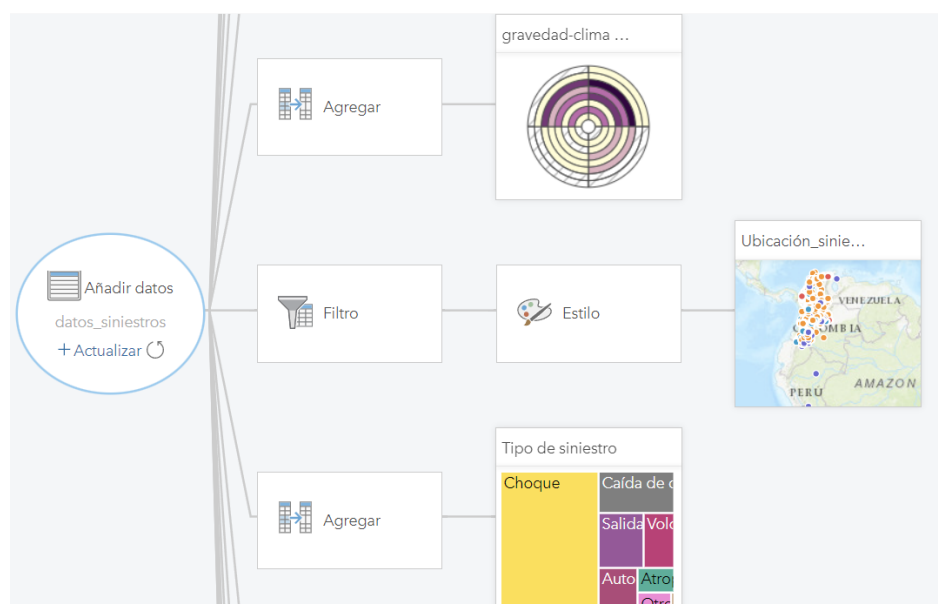


Figura 33. Vista general de todos los datos

Fuente: Propia

La figura 33, nos arroja una vista general de todas los gráficos que se tuvieron en cuenta, para el análisis, aunque en esta no se vea completo, hacia arriba y abajo existen todo lo que se usó en la investigación.

11. Conclusiones y Trabajos futuros

- El uso de ArcGIS Survey 123 es parte esencial de una investigación que necesita información de alta cantidad de datos, este permite recopilarlos de manera eficiente, sin nombrar que la opción de personalizar los formularios ayuda a capturar directamente de manera específica y precisa los datos, en este caso todo aquello que se relaciona entre sí.
- La capacidad de incorporar datos de ubicación por medio de los formularios con georreferenciación enriquece los datos en contextos muy específicos espaciales, esto permite conocer la situación presentada en las diferentes zonas del país.
- Con Insights se logró realizar análisis avanzado y exploratorio, donde se pudieron obtener tendencias en los datos y cada una de las relaciones que se generaban entre ellos, esto es tan importante a la hora de tomar decisiones y acciones ya que estas se hacen, pero informadamente, con evidencias y con datos que la sustentan.
- Se crearon visualizaciones claras e interactivas, las cuales juegan un papel muy importante a la hora de entrar a interpretar la información tanto para aquellos usuarios técnicos como para los no técnicos, datos gráficos que cualquier persona podría analizar y extraer nuevas formas de caracterizar los datos.
- Se dio cumplimiento a cada una de las funciones inculcadas en un principio por la entidad, las cuales fueron enriquecedoras para la adquisición de nuevos conocimientos y con ello para la automatización de procesos como el llevado a cabo durante la práctica.
- Se encontraron relaciones y correlaciones entre los datos, lo cual permitió establecer tendencias y estrategias de prevención en pro de la toma de decisiones, mitigando el riesgo y evitando que más eventos sigan pasando.

- Es importante conocer y disponer de conocimientos normativos vigentes, puesto que para cualquier incumplimiento ante un requerimiento proveniente de la entidad eso pasa directamente a investigación y sanciones si es el caso.
- Los informes y los gráficos visuales generados a partir de los análisis en ArcGIS Insights pueden ser compartidos con partes interesadas y así mismo pueden pertenecer a un grupo de colaboradores para entrar a manipular la información para comunicar resultados de manera clara y efectiva.
- En el proceso de pasar por una entidad para realizar la practica laboral es un crecimiento profesional muy grande, ya que se adquiere la habilidad de tomar decisiones y buscar la solución más factible a los problemas que se pueden presentar en la ejecución de las labores asignadas. Además de forjar carácter cuando a la hora de ejecutar dichas actividades se hacen bajo presión, finalmente estas particularidades forman integridad y carácter competitivo.
- En cuanto a los trabajos futuros se recomienda implementar el mismo análisis, pero para todos los programas especiales que se manejen dentro de la entidad, así mismo realizar el índice de servicio para cada uno con ayuda de Insights y con Python 3.
- También se recomienda adherir el método de regresión e interpolar los datos para obtener análisis mucho más completos, lleva un poco más de tiempo, pero lo que se obtiene aporta mucho en la manera de observar otras cosas diferentes.

12. Glosario

Back-end: Parte de sistema informático que se encarga de manejar las funciones que ocurren detrás de escena, lejos de la vista del usuario final. En otras palabras, el back end procesa la lógica, el almacenamiento y la manipulación de los datos. (Urrutia D, 2020)

Front-end: Se refiere a la parte de una aplicación o sitio web que los usuarios interactúan directamente, como la interfaz de usuario y las visualizaciones. (García I, s/f)

XLSForm: Es un estándar y una herramienta utilizada para diseñar formularios de recopilación de datos que se pueden implementar en diversas plataformas y aplicaciones, especialmente en sistemas de información geográfica (SIG). (ArcGIS Online, s/f)

PESCRI: Programa Especial Sectores Críticos de Siniestralidad, tiene el objeto de vigilar y controlar cada situación ante los eventos no deseados en las carreteras del país. (Supertransporte,2022)

JavaScript: Es un lenguaje de programación ampliamente utilizado en el desarrollo web y otras áreas de programación. (Lenguajesj, s/f)

ArcGIS Insights: Es una potente herramienta de análisis y visualización de datos geospaciales desarrollada por Esri, está diseñado para permitir a los usuarios explorar, analizar y visualizar datos de manera interactiva. (Esri, s/f)

SIG: Es una herramienta tecnológica que combina datos geospaciales (información que tiene una componente de ubicación geográfica) con tecnologías de información para gestionar, analizar y visualizar. (Ministerio de educación, s/f)

ArcGIS Survey123: Aplicación desarrollada por Esri que permite crear y gestionar formularios de recopilación de datos geospaciales de manera eficiente y efectiva. (Esri, s/f)

Correlación: Es una medida estadística que describe la relación o asociación entre dos variables. Indica la magnitud y la dirección de la relación entre estas variables. (jmp, 2020)

Dataset: Es un término utilizado en el ámbito de la ciencia de datos y la estadística para referirse a un conjunto de datos organizados y estructurados que representan información específica. (Solis, 2023)

Georreferencia: Es el proceso de asignar coordenadas geográficas (latitud y longitud) a objetos, lugares o datos para ubicarlos en la superficie de la Tierra. (Martínez D, s/f)

Software: Es un conjunto de programas, instrucciones y datos electrónicos que se utilizan para operar y controlar computadoras y otros dispositivos electrónicos.

Invias: Instituto Nacional de Vías, es una entidad gubernamental de Colombia encargada de la planificación, construcción, operación y mantenimiento de la infraestructura vial del país. (ArcGIS, s/f)

Ansv: Agencia Nacional de Seguridad Vial, es una entidad gubernamental de Colombia que tiene como objetivo principal promover la seguridad vial en el país. (Manual de estado, s/f)

CVN: Concesión vial nacional. (Lozano, 2017)

13. Referencias Bibliográficas

- Esri. (sf). Documentación de insights 3.1. Insights. https://doc.arcgis.com/es/insights/latest/get-started/pdf/insights_3_1.pdf
- Esri. (sf). Documentación de insights 3.1. Insights. https://doc.arcgis.com/es/insights/latest/get-started/pdf/insights_3_1.pdf
- Dawn Wright. (September 26, 2017). El análisis espacial relaciona el aumento de la fiebre del valle con el aumento de la temperatura del mar. Esri Blog. <https://www.esri.com/about/newsroom/blog/spatial-analysis-ties-valley-fever-increase-to-sea-temperature-rise/>
- Patricia Cummens. (22 de marzo de 2022). Mapeo del potencial de energía renovable con ayuda de NASA POWER. Esri Blog. <https://www.esri.com/about/newsroom/blog/mapping-renewable-energy-potential-nasa-power/>
- Este Geraghty, MD, MS, MPH, GISP and Ryan Lanclos. (17 de marzo de 2020). Los mapas inteligentes guían las investigaciones y acciones de COVID-19 y supervisan la eficacia. Esri Blog. <https://www.esri.com/about/newsroom/blog/gis-helps-guide-covid-19-response/>
- Md Asif Hasan Anika, James Coonsb, Nuzaira Habibb , Muhammad Ahsanul Habibc . (15 de marzo de 2023). Desarrollo de una herramienta de visualización para el modelo integrado de transporte y uso del suelo: perspectivas de un análisis de escenario de COVID-19. Science Direct. <https://www.sciencedirect.com/science/article/pii/S1877050923005574?via%3Dihub>
- Agencia Nacional de seguridad vial. ANSV - manual del estado - función pública. (s/f). Gov.Co. Recuperado el 23 de agosto de 2023, de <https://www.funcionpublica.gov.co/eva/gestornormativo/manual-estado/estructura-estado.php?id=550>
- Conceptos básicos de XLSForm. (s/f). Arcgis.com. Recuperado el 23 de agosto de 2023, de <https://doc.arcgis.com/es/survey123/desktop/create-surveys/xlsformessentials.htm>
- Correlación. (2020, agosto 7). Jmp.com. https://www.jmp.com/es_co/statistics-knowledge-portal/what-is-correlation.html
- García, I. J. B. (s/f). Backend y Frontend, ¿Qué es y cómo funcionan en la programación? Servnet.mx. Recuperado el 23 de agosto de 2023, de <https://www.servnet.mx/blog/backend-y-frontend-partes-fundamentales-de-la-programacion-de-una-aplicacion-web>

Instituto Nacional de vías. (s/f). Arcgis.com. Recuperado el 23 de agosto de 2023, de <https://inviasopendata-invias.opendata.arcgis.com/>

Lozano, C. A. M. (2017, junio 27). Delegatura de Concesiones e Infraestructura. Superintendencia de Transporte. <https://www.supertransporte.gov.co/index.php/superintendencia-delegada-de-concesiones-e-infraestructura/>

Martínez, D., & Javier, F. (s/f). Georreferenciación de documentos cartográficos para la gestión de Archivos y Cartotecas. "Propuesta Metodológica". Ign.es. Recuperado el 23 de agosto de 2023, de <https://www.ign.es/web/resources/docs/IGNCnig/CTC-Ibercarto-V-Georreferenciacion.pdf#:~:text=1.,origen%20y%20el%20sistema%20destino.>

Primeros pasos con ArcGIS Survey123. (s/f). Arcgis.com. Recuperado el 23 de agosto de 2023, de <https://learn.arcgis.com/es/projects/get-started-with-arcgis-survey123/>

¿Qué es Javascript? (s/f). Lenguajejs.com. Recuperado el 23 de agosto de 2023, de <https://lenguajejs.com/javascript/introduccion/que-es-javascript/>

Solis, D. C. (2023, mayo 1). Datasets: Qué son y cómo acceder a ellos. Openwebinars.net. <https://openwebinars.net/blog/datasets-que-son-y-como-acceder-a-ellos/>

Urrutia, D. (2020, enero 28). Qué es Backend. Arimetrics. <https://www.arimetrics.com/glosario-digital/backend>

¿Que es un SIG? -. Ministerio de Educación Nacional de Colombia: (s/f). Gov.co. Recuperado el 23 de agosto de 2023, de <https://www.mineduccion.gov.co/1621/article-190610.html>

(S/f). Wwww.um.es. Recuperado el 23 de agosto de 2023, de https://www.um.es/geograf/sigmur/temariohtml/node43_mn.html

Esri Blog. (2022, julio 7). Esri. <https://www.esri.com/about/newsroom/esri-blog/overview/>

(S/f-b). Predikdata.com. Recuperado el 23 de agosto de 2023, de <https://predikdata.com/>

14. Anexos

14.1. Presupuesto

Para poder entrar en detalle a la hora de hacer un presupuesto, primero se deben llevar a cabo la definición de los elementos que conforman el proyecto, dentro de este se encontraran aspectos como: el capital humano, materiales, equipos de tecnología, salidas de campo, la entidad financiera y las contrapartidas que se puedan generar. Por un lado, se tiene la adquisición de datos, si bien se sabe para adquirir datos se necesitan de herramientas que los proporcionen, en este caso el uso de formularios, entrevistas, informes provenientes por parte de los vigilados e inspecciones por parte de los supervisores, entre otros. También se debe contar con la limpieza y preparación de los datos, que en este caso se debe llevar a cabo el procesamiento para un análisis posterior, es decir extraer datos de prueba que posiblemente se hayan incorporado en el diligenciamiento de algún cuestionario o datos que se encuentren duplicados.

Además de ello se debe continuar con un análisis estadístico, el cual permite encontrar el objetivo general del proyecto, relaciones y correlaciones entre los datos recopilados usando el software ArcGIS Insights, lo que al momento de adquirir su utilidad poseerá un valor de carácter monetario. Seguido de ello vendría la parte de visualización de datos, en donde se presentarán los resultados y de qué manera se quieren ver, bien sea en gráficos, con mapas lo que incluye el acceso a imágenes satelitales, tablas o algún otro medio de presentación. Sin embargo, dentro de todo lo que se ha mencionado, acá viene una parte fundamental para el desarrollo de la investigación, la infraestructura tecnológica, es decir, los computadores, software, servidores y entre otras cosas que vienen regidas bajo la misma línea. Finalmente, el personal capacitado para el manejo del

software y si se quiere subir a otro nivel en asociación de datos para obtener posibles índices de servicios, se deben contar con desarrolladores analistas de datos, estadísticos y científicos de datos.

Tabla de presupuestos

Descripción	Cantidad	Precio Unitario	Precio Total
Licencia de creador en Software ArcGIS Online	1	4'000.000	4'000.000
Licencia ArcGIS Enterprise o Pro	1	300'000.000	300'000.000
Licencia ArcGIS Insights	1	50'000.000	50'000.000
Licencia ArcGIS Data Reviewer	1	17'000.000	17'000.000
Mantenimiento anual	1	45'000.000	45'000.000
Ordenador de alta capacidad	2	8'000.000	16'000.000
Desarrollador	1	7'000.000	7'000.000
Personal analista de datos	2	6'000.000	12'000.000
TOTAL			451'000.000

Tabla 1. Presupuesto

Fuente: Propia

De acuerdo con lo estipulado en la tabla 1, la entidad financiera para el desarrollo del proyecto es la Superintendencia de Transporte, es importante mencionar que el valor del programa depende de cada necesidad, ya que esta herramienta es muy amplia y por ende maneja diferentes módulos para cada utilidad, algunas de estas son licencias perpetuas, otras pueden llegar a ser de pago anualmente.

Por un lado, está la licencia del creador, es decir aquel que puede crear mapas y demás en la versión online, se paga anual y con ello viene una bolsa de créditos, esto hace referencia a que tanto se guarda en la nube de ArcGIS. Por otro lado, se tiene la licencia ArcGIS Enterprise o Pro, en este caso la que maneja la super es la básica y es perpetua, adicional a ello se debe pagar el 15%

anualmente por el mantenimiento o soporte del software. Para el caso de ArcGIS Insights se paga anualmente, y dependiendo a que otras licencias se usen, van variando los precios.

14.2. Estrategias de divulgación

Las estrategias de divulgación son realmente lo más importante después del funcionamiento del proyecto, esto se hace con el fin de dar a conocer información relevante e incluso promover el servicio que este es capaz de brindar a las diferentes entidades de acuerdo con la necesidad de cada una, se pueden difundir los resultados obtenidos en pro de dar credibilidad a todo aquel que lo quiera adquirir. Algunas estrategias fueron:

1. **Publicación en redes sociales:** Esta herramienta como bien se sabe es una de las más usadas por personas de diferentes edades, lo que intereso a muchos de ellos, por medio de esta se crearon publicaciones que además de informar promovieron el uso del servicio de manera efectiva, esto fue publicado en la pagina oficial de la superintendencia en Instagram y en la plataforma Tik-tok.
2. **Capacitaciones y/o eventos:** por medio de esta estrategia se logró realizar un evento dentro de la entidad donde junto a mi supervisor llevamos a cabo toda la preparación, aparte de esto la comunidad de Esri Colombia nos ofreció una capacitación para todas aquellas personas que quisieran saber más sobre el tema.



Figura 34. Publicidad para evento Gisday

Fuente: Propia

Con base en la figura 34, esta fue la publicidad que se hizo con respecto al evento que se llevó a cabo, el cual se tituló Gisday, donde se tenía la oportunidad de postular las iniciativas que se imaginaran en pro de la innovación y las nuevas estrategias de captar información y saberla manejar.



Figura 35. Publicación en Flickr por parte de Esri Colombia

La figura 35, muestra publicación que realizó Esri Colombia en representación del GisDay, donde junto al ingeniero Santiago acevedo se nos dio reconocimiento por las presentaciones realizadas acerca del uso del aplicativo.



Figura 36. Evento Gisday con la Supertransporte

Fuente: Propia

Como es posible observar en la figura 36, el evento fue en grande, con ayuda de la superintendencia y demás asociaciones como Esri Colombia, se tuvieron invitados a los que el tema les interesó bastante, tal como la agencia nacional de seguridad vial, el instituto nacional de vías y la superintendencia de industria y comercio.



Figura 37. Ganadores de una de las mejores propuestas futuras con el aplicativo

Fuente: Propia

Dentro del evento cualquier persona tenía la opción de proponer algo innovador que se pudiera hacer con el aplicativo y que supliera alguna necesidad de la superintendencia y junto al ingeniero Nicolas propusimos que dentro de un mapa en ArcGIS se trazaran todas las rutas nacionales de las vías a las que se le lleva control y vigilancia, para que al momento de llevar a cabo cualquier inspección por parte de la delegación se hiciera mucho más fácil, elaborada y clara la investigación.

El premio fue entregado por el jefe superintendente de la delegatura de concesiones e infraestructura, donde dio a conocer mi trabajo en la superintendencia y todas mis colaboraciones, figura 37.



Figura 38. Acompañamiento del director de concesiones e infraestructura y todo el equipo de trabajo

Fuente: Propia

Sala del evento figura 38, donde se llevaron a cabo todas las actividades, acompañamiento del jefe super intendente de la delegatura de concesiones e infraestructura y así mismo del director y demás equipo de trabajo. Allí se entablaron diversas relaciones con distintas entidades, propósitos, colegas. Se divulgó información de interés y se debatió sobre lo que se tiene y como se puede seguir mejorando con aquellos que comparten el mismo enfoque.