



Vigilada Mineducación

USO DE DATOS ABIERTOS PARA LA IDENTIFICACIÓN DE POBLACIONES
VULNERABLES ANTE EVENTOS DE EMERGENCIA EN LA CIUDAD DE
MEDELLÍN

Use of open data for the identification of vulnerable populations in emergency
event in the city of Medellín

SEBASTIÁN OSPINA PORRAS

Trabajo de Grado

Asesor

Juan David Ospina Arango

UNIVERSIDAD EAFIT
ESCUELA DE INGENIERÍAS
MAESTRÍA EN CIENCIAS DE LOS DATOS Y LA ANALÍTICA
MEDELLIN

2021

Agradecimientos

Agradezco a mi tutor Juan David Ospina Arango por el gran acompañamiento brindado durante todo el desarrollo del presente documento.

Al profesor Juan Camilo Sánchez de la Universidad de Antioquia por su valiosa ayuda para entender el contexto de la logística humanitaria.

A mi familia y mis compañeros de estudio que estuvieron acompañándome y apoyándome en mi formación académica.

Tabla de contenido

Agradecimientos	2
1 Resumen	7
2 Planteamiento del Problema	10
3 Justificación.....	11
4 Objetivos	14
4.1 Objetivo general:.....	14
4.2 Objetivos específicos:.....	14
5 Estado del arte y Marco teórico.....	14
5.1 Logística humanitaria:.....	14
5.2 Logística humanitaria y la tecnología:.....	18
5.3 Logística humanitaria y sistemas de información geográfica (SIG):	19
5.4 Logística humanitaria ante eventos naturales:.....	21
5.5 Logística humanitaria y ciencia de datos:	22
5.6 Gestión del territorio y Sistemas de Información Geográfica:	23
6 Metodología.....	25
6.1 Identificación del escenario y principios.....	25
6.2 Fuente de datos.....	26
6.3 Análisis fuentes de datos disponibles	27

6.3.1	Datos Nacionales	27
6.3.2	Datos Municipales	28
6.4	Modelo de datos	30
6.4.1	Modelo de datos relacional	30
6.4.2	Modelo de datos geoespacial:.....	31
6.5	Explotación de los datos y construcción del modelo.....	32
6.5.1	Lectura de los datos	32
6.5.2	Criterios de limpieza de datos	34
6.5.3	Transformación de los datos	35
6.5.4	Consolidación de los datos por modelo relacional	36
6.5.5	Consolidación de los datos por modelo geoespacial.....	37
6.5.6	Selección de variables finales	38
6.5.7	Modelo de consulta y generación de resultados	38
6.6	Arquitectura de datos en producción	39
6.7	Visualización de los resultados.....	40
7	Casos de estudio.....	41
7.1	Planteamiento de Hipótesis	41
7.2	Caso Inundación.....	43
7.3	Caso Deslizamiento.....	51

7.4	Caso avenida torrencial:	57
8	Discusión y Conclusiones.....	64
9	Trabajos futuros	65
10	Aspectos éticos	66
11	Uso de los datos	66
11.1	Políticas de uso	66
11.2	Archivación de datos y modelo	66
12	Bibliografía	67
	Figura 1. Arquitectura de datos en producción	40
	Figura 4. características de las personas de la zona 1	45
	Figura 5. características de las personas de la zona 2	45
	Figura 6. características vivienda de la población de la zona 1	47
	Figura 7 características vivienda de la población de la zona 2	48
	Figura 8. características del entorno de la zona 1	49
	Figura 9. características del entorno de la zona 2	50
	Figura 10. Características de la población de la zona 3	52
	Figura 11. Características de la población de la zona 4	52
	Figura 12. características de las viviendas de la zona 3.....	54
	Figura 13.características de las viviendas de la zona 4.....	54
	Figura 14. características del entorno de la población de la zona 3	56

Figura 15.características del entorno de la población de la zona 4	56
Figura 16.Características de la población de la zona 5	59
Figura 17.Características de la población de la zona 6	59
Figura 18. Características de la vivienda de la zona 5.....	61
Figura 19.Características de la vivienda de la zona 6.....	61
Figura 20.Características del entorno de la zona 5.....	63
Figura 21.Características del entorno de la zona 6.....	63
Tabla 1.Fuente de los datos del modelo relacional.....	31
Tabla 2. Descripción de los datos cargados en la etapa de consolidación	33
Tabla 3. Valores P prueba Chi ²	42
Tabla 4. Zonas seleccionadas para el caso de estudio	43
Tabla 5. Zonas seleccionada caso deslizamiento.....	51
Tabla 6. Zonas seleccionada caso avenida torrencial	57

1 Resumen

Cuando un desastre natural ocurre, y este genera daños a las poblaciones del área de influencia del evento, existe la necesidad de realizar actividades de logística humanitaria, rápidas y oportunas que permita la adecuada atención de las comunidades, frente a este contexto los datos se convierten en un factor importante ya que estos nos permiten tomar decisiones con menor nivel de incertidumbre. Países como Japón, han implementado el uso de estos para la preparación, planificación y gestión ante desastres naturales como terremotos o tsunamis, han buscado generar capacidades resilientes en las poblaciones (Banco Mundial, 2020).

En el caso colombiano hay datos que se podrían utilizar ante estas situaciones, específicamente en la atención de desastres ocurridas por eventos naturales, dado que cerca del 24% de la población de la ciudad de Medellín habita en zonas de alto riesgo (Alcaldía de Medellín, 2013) y dicha exposición puede detonar una emergencia. La atención a estos eventos, adicional a la financiación y obtención de recursos, requiere de planeación logística, y una de las formas para disminuir la incertidumbre en la toma de decisiones durante esta etapa es el uso de los datos.

Dado que la atención a desastres requiere de acciones rápidas y oportunas, el uso de datos que faciliten la identificación y caracterización de la población afectada se convierte en un medio adecuado en estas situaciones. Para lograr esto se hace necesario contar con información disponible de la zona y su población. Existen fuentes de información dispersa y con distintos niveles de detalle, por lo que agregar convenientemente la información es necesario para llegar a resultados relevantes.

Adicional a lo anterior, es posible que para la mayoría del público no sea suficiente solo con contar con la información consolidada, sino que se requiere que la información se pueda resumir o agregar de manera lo que demanda tener un proceso de analítica descriptiva que entregue información relevante al público objetivo.

Ante la problemática expuesta, de ausencia de herramientas o métodos que soporten la toma de decisiones en un menor tiempo, se propuso desarrollar un modelo que, ante la ocurrencia de un desastre, permita tener una caracterización de la población afectada y su entorno. Para lograr esto, en primera instancia se procedió a seleccionar un área de estudio, y dado que cada población o subregión del país presenta características sociodemográficas y culturales diferentes, se realizó un análisis de la información disponible, en donde se decidió desarrollar un modelo en el contexto de la ciudad de Medellín por sus progresos en disponibilización de datos abiertos.

Una vez seleccionada el área de estudio, “ de desastre posibles para la ciudad para establecer casos de uso. A partir de información del departamento de gestión de desastres de la ciudad en donde se logró identificar que los eventos de inundación, avenidas torrenciales y deslizamiento eran los escenarios con mayor probabilidad de ocurrencia en el área.

Con respecto a los datos necesarios para la construcción del modelo se hizo uso de los datos abiertos que se encuentran disponibles en la página nacional de datos abiertos y en el portal de datos abiertos del municipio de Medellín.

Durante la construcción del modelo se realizaron procesos de transformación y consolidación de la información. El primer proceso consistió en la selección de datos referentes a la ciudad de Medellín, y en la categorización de las variables que se encontraban con valores numéricos. El segundo proceso consistió en consolidar la información mediante metodologías relacionales y geoespaciales que permitieron tener todos los datos de la población y el entorno bajo una misma base consolidada. Una vez finalizados los dos procesos anteriores, se construyó el modelo de consulta, el cual, a partir de un área ingresada, este genera unos datos de salida de la población contenida en esta. Para explotar el modelo de datos resultante se propone un tablero de Power BI donde se pueda visualizar un resumen que caracteriza la población, su vivienda y su entorno para tener un contexto de ellos, un dimensionamiento de las ayudas necesarias, y un precenso de la misma.

El tablero de Power Bi anteriormente mencionado fue desarrollado para facilitar la lectura de los datos y su correspondiente análisis. Consta de 3 páginas que cuentan una historia en donde primero se caracteriza la población, luego se describe las condiciones de sus viviendas, y por último su exposición a temas de seguridad, buscando que se permitan generar estrategias de acompañamiento acordes a sus necesidades.

Para mostrar cómo se puede usar el modelo se plantearon tres escenarios de desastre (Inundación, Avenida torrencial y Deslizamiento) en distintos lugares de la ciudad y se muestra que existen diferencias en las necesidades de la población según su área de influencia, para cada escenario se seleccionaron por ubicación

dos tipos de poblaciones: la primera localizada en barrios de estrato bajo, y la segunda en barrios de estrato alto. Cuando se ejecutaron los escenarios, se analizaron los resultados obtenidos, en donde se logró encontrar que si existían diferencias entre ambas poblaciones.

Frente a los resultados obtenidos en los escenarios, se logra identificar la capacidad que tiene el modelo desarrollado de cuantificar y caracterizar a una población afectada en términos del individuo, su vivienda y su entorno, generando así información relevante en el diseño de un plan de logística humanitaria, ya que este aporta una aproximación a la etapa de censo de personas y viviendas afectadas, y a su vez a la implementación de estrategias de atención y acompañamiento de la población.

En próximos pasos se recomienda robustecer la información social y del entorno, porque a la fecha si bien existe esta información a nivel de comuna mediante la encuesta de calidad de vida, es necesario llevar u obtener esta información a un nivel de granularidad mayor. También se propone escalar el modelo a otras regiones, o ampliarlo a toda el área metropolitana del valle de aburra, ya que a medida que el territorio crece las fronteras entre municipios son cada vez menos claras, por lo que se hace necesario poder obtener una mayor área de estudio.

2 Planteamiento del Problema

La atención a poblaciones frente a afectaciones dadas por eventos naturales catastróficos, o humanitarios, se facilita mediante una planeación logística, adicional a los recursos obtenidos para las poblaciones. Ante lo anterior, existen retos de la logística humanitaria que requieren el apoyo de datos y de metodologías en toma de decisiones para realizar dicha planeación.

Uno de los retos de la logística humanitaria en cualquier contexto es el poder suplir las ausencias de conocimiento sobre las poblaciones afectadas y su entorno en la etapa de planeación, en dónde se buscar tener respuestas sobre: ¿Qué se necesita?, ¿Cuándo se necesita?, ¿Cuándo se necesita?, y ¿Adonde se necesita? (OPS, 2000); ya que pueden quedar poblaciones por fuera de la ayuda. Por último, otro reto es el desconocimiento de las condiciones sociales, y de seguridad en dichas áreas, generando puntos ciegos en la observación del entorno.

Ante las necesidades expuestas, el uso de datos públicos y disponibles de las poblaciones y su entorno se convierte en un recurso para la construcción de modelos, o metodologías que ayuden en este caso a solucionar los vacíos en el conocimiento del entorno de las poblaciones para la atención y/o preparación a eventos naturales.

3 Justificación

El modelo o trabajo propuesto busca tener importancia en las etapas de prevención, preparación y planeación de la logística humanitaria, ya que aporta conocimiento y entendimiento de poblaciones o zonas de influencia, cubriendo así vacíos o necesidades que se tienen en temas sociodemográficos, del entorno y localización identificados en la literatura (Cinnamon, 2020).

Se puede observar en el contexto colombiano la importancia de este tipo de trabajos, ya que en el país el porcentaje de destrucción de viviendas por inundación y deslizamiento representan el 43% y 10% respectivamente, y el 10% y 36% de las muertes ocurridas en desastres naturales (Banco Mundial, 2012). Ante estas cifras se requiere de mayor número de herramientas que permitan robustecer el modelo de atención y gestión a desastres, como por ejemplo la caracterización previa de las personas y su entorno, generando acciones oportunas y certeras.

En este tipo de situaciones, la información geoespacial ha sido utilizada como herramienta de observación del entorno, debido a que esta tiene la facilidad de representar información en los mapas en tres niveles geométricos: poligonal, lineal, y puntual. En el nivel poligonal nos podemos encontrar con información como manchas de inundación, áreas de influencia, entre otros; al nivel lineal se han representado rutas, o líneas de carretera; y finalmente a nivel puntual se puede encontrar información de ubicación exacta de las personas, sitios de interés, etc.

Usualmente obtener este tipo de información geoespacial presenta sus retos, entre los que están el acceso y la calidad de la información, pero estos retos han sido superados de manera gradual debido a las iniciativas públicas de crear bancos de datos abiertos. Estos son ambientes repositorios de información pública que se ha generado a partir de la recopilación de datos en proyectos, o en el día a día de las entidades implicadas, y que se encuentran a disposición de la población para su uso y beneficio. El acceso a esta información también ayuda a superar el reto de información en la logística humanitaria.

Ante lo anteriormente expuesto, para la construcción del modelo se contó con información geoespacial de las personas y de algunas variables sociodemográficas que para un caso u otro son relevantes y dan soporte al modelo, permitiendo obtener estimaciones de la cantidad de afectados, y sus características sociales y del entorno con mayor granularidad. Esto, disminuirá la incertidumbre que en ocasiones se genera por solo tener la información a nivel comunal o barrial (nivel poligonal), y que generaliza la información a nivel de su área, perdiendo relevancia a la hora de tomar decisiones.

El propósito de este trabajo es construir un modelo de identificación de poblaciones, que permita ayudar a los involucrados en logística humanitaria en la planeación a la

atención de las poblaciones más expuestas, y/o con mayor vulnerabilidad social y económica. Buscando así que las medidas y estrategias que se tomen sean con conocimiento, generando un impacto social y un verdadero cambio en el área o zona afectada.

4 Objetivos

Los objetivos del presente trabajo son:

4.1 Objetivo general:

Construir un modelo de identificación geoespacial de poblaciones basado en datos abiertos del gobierno nacional y local, como herramienta de apoyo en la toma de decisiones ante eventos de emergencia en la ciudad de Medellín.

4.2 Objetivos específicos:

- Desarrollar un modelo de datos relacional unificado a partir de las fuentes de información geoespacial de datos abiertos.
- Proponer un algoritmo de agrupación que permita el dimensionamiento y caracterización de la población según el tipo de evento y área de influencia.
- Demostrar el uso del modelo en un escenario hipotético de materialización de un evento natural en la ciudad

5 Estado del arte y Marco teórico

5.1 Logística humanitaria:

La logística humanitaria es una herramienta de apoyo que busca tener una adecuada planeación para la atención y acompañamiento a las poblaciones. A partir de la literatura recopilada, se ha observado que este tema ha sido en su mayoría abordado en el contexto de desastres de la naturaleza o de guerra, pero a pesar de que su desarrollo ha sido en dichos temas, esta posee elementos y fundamentales importantes destacados por los autores que se mencionaran más adelante, que pueden ser escalables en un contexto del flujo de ayudas a las poblaciones de bajo recursos o menos favorecidas. Hablando de las características fundamentales de la logística humanitaria, Osorio (2016) menciona que en esta es importante identificar las necesidades de la población, como también la de los involucrados en la cadena de suministro, y a su vez conocer si los recursos disponibles satisfacen dichas necesidades, y por último identificar factores relevantes en el entorno de la zona. Para el autor el desconocimiento de estas variables antes de la llegada de recursos a una población o área puede generar resultados no esperados tanto en la población, como en los involucrados, y también en una mala logística para la entrega de recursos que se pueden asignar a las personas que no son el objetivo de la iniciativa.

Analizando las diferentes maneras que se puede abordar el problema, Osorio (2016), realiza una caracterización previa de los modelos cuantitativos existentes para justificar su implementación en modelos estocásticos (identificado por el autor como usado en algunos estudios en el tema), dándole relevancia a la identificación de estudios previos cuyo enfoque principal ha sido sobre temas de investigación de

operaciones en la optimización tanto de recursos como del tiempo de respuesta a eventos (la investigación está enfocada a atención de desastres), implicando un abordaje del problema con elementos de probabilidad y simulación de escenarios, buscando que estos últimos sean los más realistas posibles. Finalmente realiza una evaluación de la metodología en casos reales ocurridos en Colombia, para ajustar el modelo, buscando que este ayude a mejorar la atención y respuesta en temas de logística humanitaria.

Siguiendo en el contexto colombiano, Gómez, et al (2017) identifican y analizan una red de ayuda humanitaria a partir de un caso de estudio de la ciudad de Manizales, ya que para los autores este es un factor relevante en la logística humanitaria, debido a que la coordinación entre diferentes actores es uno de los mayores problemas cuando ocurre un desastre. Para abordar dicha problemática realizaron una identificación de los actores, sus vínculos entre sí, el peso de estos, y mediante grafos realizaron el análisis de la red, en donde se encontró las deficiencias y brechas que a la fecha del estudio se presentaban en la red de logística humanitaria, permitiéndoles así, generar recomendaciones aplicados a la ciudad, y finalmente recomendando realizar estudios enfocados en el refinamiento de redes.

Continuando con el análisis de como se ha abordado las diferentes problemáticas entorno a la logística humanitaria, se encuentra la aplicación de las metodologías de lógica difusa y Proceso Jerárquico de Redes (AHP) realizado por Galleguillos-Pozo et al (2019), mediante estas los autores buscaron seleccionar los centros de acopio apropiados ante la ocurrencia de un desastre, a través de un caso de estudio

en Ecuador, e implementando valoraciones a las opiniones de los decisores o actores buscaron definir las ubicaciones más adecuadas, debido a que es importante identificar dichas zonas seguras como medida de precaución a la ocurrencia de un nuevo evento en el corto plazo.

Otra forma de abordar la logística humanitaria se ha identificado en Ochoa et al (2017), donde hablan en su investigación sobre la importancia que tiene la diversidad cultural, y las condiciones de ocurrencia de un evento en el comportamiento de las personas ante este, a través de escenarios y comportamientos simulados generan modelos de apoyo a la toma de decisiones ante la respuesta en dichas situaciones. Adicional se destaca en dicho estudio el uso de datos para simular comportamientos sociales, dado que es una rama que viene siendo explotada poco a poco en búsqueda de entender por qué las personas actúan de cierta forma en ciertos momentos y más ante eventos que necesitan respuesta rápida.

Por otro lado, abordando la ubicación de puntos estratégicos en logística humanitaria como otro factor relevante pero complejo determinar en esta, Abazari et al (2020) desarrollan un modelo multiobjetivo con parámetros inciertos en eventos de catástrofe, en donde consideran factores como los centros de socorro, las áreas de demanda, los medios de transporte, y el tipo de insumo, para determinar puntos en una zona antes del desastre, para cuando ocurra el evento a partir del modelo se determine la cantidad de recursos necesarios y como hacerlos llegar.

Por último, Jabbour et al (2019) hacen un exhaustivo análisis de la literatura sobre logística humanitaria, y la administración de la cadena de distribución, considerando factores como el contexto económico del país de ocurrencia, el enfoque, la metodología utilizada, el tipo y la fase del desastre, actores, y región o ubicación de los autores, logrando así sintetizar y resumir varios estudios sobre el tema, encontrando que la mayoría de ellos estaban basados principalmente en el componente teórico y algunos en configuración de modelos.

5.2 Logística humanitaria y la tecnología:

Si bien los modelos que se han desarrollado para dar solución a los retos que se presentan en la logística humanitaria son relevantes, es importante llevar estos modelos a disposición de la sociedad a través de la tecnología que se ha convertido en un habilitador de la divulgación del conocimiento y la información. Por ejemplo Crumbly & Carter (2015) realizaron un análisis del impacto de las redes sociales en el desarrollo de nuevos servicios en logística humanitaria, haciendo mención de obras benéficas para la recolección de productos de alimento y otros insumos necesarios para las poblaciones empobrecidas, en donde el principal reto en estas organizaciones es lidiar con las fechas de caducidad de algunos productos, para que estos lleguen antes de dicha fecha, es por eso que exploraron como mejorar la eficiencia y eficacia operativa a través de la tecnología para el seguimiento a las tareas, e incluyendo a las redes sociales como soporte a la gestión de la cadena de suministro.

Por otro lado, Abushaikha & Schumann-Bölsche (2016) evalúan como las tecnologías implementadas en teléfonos móviles puede apoyar las operaciones en logística humanitaria, y en cuales situaciones podrían ser idóneas o no en este ámbito buscando proporcionar información relevante en la evolución y mejora de la metodología. Para su análisis hicieron uso de casos de refugiados en África y Europa, en donde concluyen que los teléfonos móviles pueden aumentar la eficiencia y efectividad de la logística humanitaria en un contexto de refugiados, y posiblemente en otros contextos, pero aclaran que cada desarrollo debe responder a las necesidades específicas de la población a impactar y hacen énfasis en una frase: “soluciones que están salvando vidas en la respuesta humanitaria”.

5.3 Logística humanitaria y sistemas de información geográfica (SIG):

Dado que el modelo a desarrollar se hará a nivel geoespacial, es importante resaltar como los sistemas de información geográfica han sido utilizados en respuesta a las necesidades de la logística humanitaria. Cinnamon (2020) exploró el mapeo humanitario como un adicional a la gestión de información humanitaria, mencionando la relevancia de tener información confiable de la población, el entorno, y las condiciones actuales de la zona para un buen desarrollo de la herramienta, es por esto que los mapas toman relevancia en la planificación por su facilidad de visualizar el entorno, pero para el autor aún existe el reto de que una región posea una infraestructura de datos de calidad e incluso que posea datos relevantes para llevarlos a estos niveles espaciales. Por otro lado reconoce como la

tecnología ha promovido que los sistemas de información geográfica en el ámbito humanitario este en crecimiento, especialmente en el ámbito público, produciendo así nuevos datos y mapas en la generación de ayudas humanitarias, dado que se requiere conocer las ubicaciones exactas de las zonas afectadas en atención de los desastres, y principalmente cuando existen poblaciones que no tienen conocimiento de la zona del evento ocurrido y quieren ayudar, o también la aplicación del seguimiento en tiempo real, en donde se pueda monitorear la situación de la zona. Por último, el autor identifica algunos campos de acción a explorar en este ámbito, y es como la inteligencia artificial, y los sistemas de información geográfica permite mejorar la calidad de los datos y en agilizar la toma de decisiones basada en evidencia.

Por otro lado, Ortiz (2020) desarrolló un metaanálisis sobre como los sistemas de información geográfica (SIG) se han convertido en una herramienta importante en la atención a desastres ocasionados por eventos de la naturaleza, ya que ha promovido el aumento en el impacto positivo de las ayudas a las comunidades o poblaciones. Adicional a lo anterior, identifica que la mayoría de aplicaciones de los SIG se ha llevado en temas de análisis de redes y sistemas de distribución, cuando estos pueden abarcar más temas o retos generados en el contexto de logística humanitaria, por lo que el autor por medio de su metaanálisis ejemplifica algunas aplicaciones adicionales en los SIG, señalando las fortalezas que esta herramienta posee, en donde una de las ramas que presenta aplicación es la involucración de

información sociodemográfica y de riesgos en los análisis espaciales para logística humanitaria.

Los sistemas de información geográfica también pueden ser un soporte en el tema de distribución de recursos, Qureshi & Taniguchi (2020) desarrollaron un modelo de logística humanitaria considerando los recursos disponibles, este surge de la necesidad de tener una adecuada planificación ante eventos catastróficos. Para la construcción del modelo, los autores realizaron un estudio de caso de terremoto en la ciudad de Osaka, buscando la ubicación eficiente del depósito de recursos, y cuál sería la estrategia de suministro a través de ruteo más óptimo, ya que, en eventos catastróficos, el tiempo de respuesta es importante

5.4 Logística humanitaria ante eventos naturales:

Debido a que se realizará el estudio bajo uno de los escenarios de desastre del municipio de Medellín, y buscando aplicaciones frente algunos eventos, se ha encontrado que algunos autores como Díaz-Delgado & Iniestra (2014) desarrollaron un proceso de diseño de operación de logística humanitaria ante eventos de emergencia ocasionados por inundación. Para su análisis, primero buscaron generar escenarios hidrológicos de lluvia para generar inundaciones en un área en específico que permita determinar la magnitud del evento, una vez conocidos estos resultados, generan algunas propuestas en torno a la atención del evento, y es validado posteriormente a partir de un ejemplo, en donde encuentran la importancia de estar siempre preparados ante estos eventos.

Observando la importancia de la colaboración entre diferentes actores, pero identificando problemas en la optimización de los recursos provenientes de estos que tienden a producir una escasez de los mismos, Rodríguez-Espíndola, et al (2018) desarrollan un sistema de preparación de desastres combinando sistemas de información geográfica, y optimización multiobjetivo para la toma de decisiones. A partir de estas dos metodologías buscan cual es la ubicación adecuada de los almacenamientos, considerando que no esté en zonas inundables, pero que también considere el número de personas necesarias para la operación. Para llevar el modelo a situaciones reales, se ejemplificó bajo un escenario de inundación ocurrido en México, en donde muchas entidades participaron, pero al no saberse manejar se generó un alto costo sin lograr los niveles de satisfacción posibles, por lo que al llevarse a la metodología propuesta se observó una mejoría en el desempeño de las operaciones en logística humanitaria para atender el evento.

5.5 Logística humanitaria y ciencia de datos:

Abordando con mayor detalle cómo la ciencia de los datos más allá de la tecnología puede ser relevante en dar solución a algunas situaciones en logística humanitaria, Gupta, et al (2019) emplearon una búsqueda exhaustiva de aplicaciones del análisis de datos en la preparación o atención a desastres, encontrando que por ejemplo la ciencia de los datos aplicada en datos no estructurados como los obtenidos en redes sociales pueden ayudar en la ejecución de operaciones de ayuda humanitaria.

Por otro lado, y buscando una solución práctica en estudios de caso, Tacheva & Simpson (2019) realizan un análisis de redes humanitarias en la logística

humanitaria a partir de dos casos de estudio, utilizando datos secundarios y análisis de gráficos para lograr identificar la utilidad de la metodología y también los motivos por los cuales las redes participan a la atención de este tipo de eventos. A partir de esto logran encontrar que el análisis de redes humanitarias facilita el entendimiento de los patrones de estructura comunitaria y también a identificar las motivaciones de cada agrupación.

5.6 Gestión del territorio y Sistemas de Información Geográfica:

Dado que uno de los objetivos del modelo es clasificar o agrupar las poblaciones, se hace importante evaluar las aplicaciones que se han realizado a través de Sistemas de Información Geográfica (SIG) en la gestión o clasificación de poblaciones en un territorio. Kilicoglu et al (2021) dadas las evidencias de crecimiento poblacional, y la migración de las poblaciones rurales a las grandes ciudades, identifican uno de los mayores problemas en temas de densidad poblacional el cual es la ubicación estratégica de los nuevos desarrollos urbanos y habitacionales desarrolla un modelo de toma de decisiones multicriterio integrando herramientas SIG en una población de Turquía. Este modelo propuesto evalúa el nivel de exposición a riesgos naturales, como también a riesgos físicos como líneas de tensión de energía e igualmente la biocomodidad que es un indicador elemental en temas de valoraciones de calidad de vida.

Por otro lado Dinh y Pham (2020) a través de los Sistemas de Información Geográfica y Teledetección cuantifican el impacto de la urbanización sobre una ciudad patrimonio cultural a nivel mundial de Vietnam encontrando una relación

positiva entre el aumento de la población, y el turismo con la expansión del suelo urbano, lo que genera un deterioro del patrimonio cultural, lo que en un primer acercamiento concluyen que estas tendencias pueden afectar la calidad de vida de la ciudad en el largo plazo.

Uno de los temas relevantes en las poblaciones en la medición de calidad de vida es la calidad del aire, y el monitoreo de este en algunas ocasiones se representa a nivel de mapas, Oji y Adamu (2021) evalúan los datos de calidad del aire de algunos sensores del área metropolitana de Kano – Nigeria para crear mapas de distribución de la calidad del aire en la zona de estudio, a partir de estos resultados calculan un indicador de contaminación encontraron que aquellas poblaciones cercanas a zonas industriales y comerciales encontrando así que la calidad del aire esta correlacionada con el uso del suelo, por lo que es un insumo en el manejo de poblaciones y gestión del territorio.

Con el surgimiento del COVID-19 surgieron amenazas para la salud y los sistemas socioeconómicos, Kanga et al (2021) a partir de esa afirmación proponen un marco de evaluación de riesgo del COVID-19 como una herramienta facilitadora en la toma de decisiones en donde se involucra amenaza y vulnerabilidad asociados a la pandemia. El desarrollo de este marco se hace sobre una ciudad de India se realizaron hipótesis de índice de riesgo del covid y teniendo en cuenta variables como la densidad de la población, disponibilidad de recursos limpios y ubicación que se encuentran relacionados con el contagio de la enfermedad se desarrolla un modelo multicriterio a partir de los SIG encontrando en los resultados zonas donde

existía mayor riesgo para la población por lo que propusieron cuales zonas son las que se deben priorizar para dar una buena gestión de la crisis de la pandemia.

Finalmente y con mayor relación a los riesgos de la naturaleza, Ranman et al (2021) analizan la relación de las inundaciones con el cambio en el uso de la tierra, el crecimiento de la población y de las carreteras en la ciudad de Bangladesh, a través de modelos de machine learning como redes neuronales, arboles de clasificación, entre otros propuestos por los autores, y combinando dichas metodologías con el procesamiento de mapas en herramientas SIG, analizan la cobertura terrestre del área de estudio entre los años 2000 y 2017. Uno de los hallazgos del análisis fue que durante ese periodo las áreas con mayor susceptibilidad o riesgo a inundación habían aumentado aproximadamente 20%. Los autores encuentran en sus hallazgos gran aporte en la formulación de políticas y en la gestión de riesgos y recursos.

6 Metodología

Para lograr identificar y clasificar las poblaciones vulnerables o afectadas ante eventos de la naturaleza o humanitarios a nivel geoespacial se hizo necesario de pasos previos que van desde la recolección de la información, hasta la construcción del modelo de consulta y visualización de los resultados.

6.1 Identificación del escenario y principios

Como el modelo propuesto está planteado para eventos de desastres naturales, se hace necesario contar con escenarios reales o hipotéticos para validar su utilidad frente a estas situaciones, por lo que se formularon tres casos de uso a partir del

planteamiento de tres escenarios: Inundación, Deslizamiento y Avenida Torrencial. Estos escenarios fueron identificados a partir de información expuesta por el Departamento Administrativo de Gestión del Riesgo de Desastres – DAGRD de la ciudad de Medellín, en donde se expone que para la ciudad estos eventos son los de mayor riesgo para algunas zonas de la ciudad, ya sea por la ocurrencia pasada de estos en la zona o por estudios de campo previos. Con el insumo resultante de la identificación anterior, las necesidades de la logística humanitaria y la información disponible de los datos abiertos se realizó el análisis de los datos a partir de cada uno de los escenarios.

6.2 Fuente de datos

Las fuentes de datos utilizadas fueron obtenidas a partir de los sistemas de datos abiertos a nivel nacional y municipal, las cuales posteriormente recibieron un tratamiento de los datos que se describirá posteriormente, junto con las consideraciones tomadas en el presente trabajo. A continuación, se presentan las fuentes de datos utilizadas y su aporte dentro de la construcción del modelo:

- [Censo poblacional 2018 - DANE](#): Caracterización de población
- [Mapa vulnerabilidad - DANE](#): Caracterización de población
- [Mapa amenazas naturales - DAGRD](#): Caracterización del entorno
- [Encuesta calidad de vida](#): Caracterización de población
- [Datos de seguridad](#): Caracterización del entorno

Los conjuntos de datos anteriormente mencionados fueron seleccionados en primera instancia por el fácil acceso a estos, ya que su uso es libre; otra de las razones se

debe a que los datos del censo poblacional facilitan un entendimiento de las poblaciones y su situación económica y social en el momento de la muestra, y también por que poseen un nivel de ubicación geoespacial a nivel de manzana predial, lo que permite hacer análisis con mayor precisión. Por otro lado, los datos de seguridad y de riesgos de la naturaleza también se encuentran en la misma escala que los datos del DANE por lo que ayuda a entender el entorno de las poblaciones, y a su vez facilita los análisis y estimaciones con menor incertidumbre a partir de escenarios catastróficos.

6.3 Análisis fuentes de datos disponibles

El modelo propuesto se desarrolló a partir de datos abiertos que se pueden encontrar con cobertura nacional, regional, o municipal. Según la disponibilidad de los datos, y la relevancia de la información que estas pueden aportar, se delimitó la cobertura y escala a la cual el modelo se construirá, frente a la necesidad de contar con información a un nivel de detalle cercano al individuo por individuo, se analizó las diferentes fuentes de datos, con la información que esta podía suministrar y a que cobertura, empezando desde la escala nacional, hasta la municipal.

6.3.1 Datos Nacionales

Cuando se exploró la información a nivel nacional que podrían aportar al modelo se identificaron tres fuentes: [Cifras consolidadas](#) del censo 2018 a nivel de manzana del DANE, Datos del [mapa de vulnerabilidad](#) a nivel de manzana del DANE, y [Microdatos](#) de censo 2018 del DANE.

Los datos consolidados del censo 2018, contiene información a nivel de manzana, en donde cada manzana tiene sus coordenadas, y algunas características de esta como cuantas personas habitan en ella, cuantos hombres, cuantas mujeres, entre otras. Mas allá de la información que contiene este conjunto de datos, su funcionalidad y participación en el modelo está atado a la intercepción geoespacial entre el polígono del área de afectación con los polígonos de manzanas que contienen un código único que facilita la selección de las personas en dicha área.

Por otro lado, el mapa de vulnerabilidad del DANE, el cual fue construido a partir de la información del censo 2018 junto con información de salud, su información es relevante debido a que esta cuenta con el Índice de Pobreza Multidimensional (IPM), en donde se calculó el índice para cada manzana, lo que permite un entendimiento de la calidad de vida de las personas.

Finalmente, el censo poblacional 2018 del DANE cuenta con microdatos, los cuales se encuentran a un mayor nivel de detalle para el análisis de información, ya que trae consigo las respuestas a cada una de las preguntas realizadas a nivel de vivienda, y población, por lo que se puede trabajar con esta información y llevarla a algo geoespacial, aunque estos microdatos no poseen coordenadas de la manzana, si tiene el código de la misma de cada una de las familias encuestadas, lo que se pudo realizar una unión con la información geoespacial consolidada del censo que cuenta con las coordenadas y tiene el código de la manzana.

6.3.2 Datos Municipales

Mirando los datos en una cobertura más pequeña, se encontró que solo los municipios de Medellín, y Sabaneta contaban con información que puede ser útil en el desarrollo del modelo, pero frente a las necesidades del modelo desde la caracterización de la población, hasta la caracterización del entorno, el municipio de Sabaneta solo contaba con información de la población, por lo que se decidió utilizar la información disponible en el municipio de Medellín.

A nivel de barrios, se cuenta con la información recolectada en la encuesta de calidad de vida del municipio, está a parte de valorar las condiciones de vida de las personas también evalúa la percepción que tiene la población frente algunos temas como salud y educación. Esta información sirvió como un complemento a la información del DANE, primero porque el número de encuestados es menor al de la población del DANE, y la información está a nivel de barrio por lo que no era posible identificar con mayor precisión la ubicación de las poblaciones con mayores necesidades, sirviendo entonces como un elemento de contraste al resultado obtenido del DANE.

Una parte de la información de delitos del municipio, también se encuentra a nivel barrial, lo que permitió identificar algunos tipos de criminalidad en las zonas, este tipo de información aporta elementos en la toma de decisiones en la planeación logística de atención a desastres, ya que conocer las condiciones de seguridad de la zona como temas de hurto, falsificación y corrupción ayuda a prevenir los riesgos asociados a estos delitos (Khan, Lee, & Bae, 2019).

Ya en una granularidad de los datos más fina, se tiene a nivel de coordenadas información relevante sobre el entorno de las personas: mapa de amenazas a riesgos de la naturaleza lo que permitió establecer escenarios; información de violencia intrafamiliar, secuestro, homicidio, extorsión, convivencia, para identificar el entorno de las poblaciones. Por otro lado, el municipio cuenta con un sistema de alerta temprana de casos sociales de Medellín, donde se identifica las zonas identificadas con ciertas problemáticas sociales y el nivel que esta presenta.

Por otro lado, en este mismo nivel de coordenadas, se cuenta con información de espacio público, que en temas de logística humanitaria es útil para identificar zonas de encuentro donde la gente pueda asistir más fácilmente y acudir por las ayudas a entregas.

6.4 Modelo de datos

Dado que los datos provienen de diferentes fuentes y con diferente estructura, se hizo necesario diseñar un modelo de datos que permitiera consolidar los datos y también la ejecución del modelo en producción, para esto se hizo necesario realizar un proceso de unión de datos a partir de llaves o columnas en común, y también un proceso de intersección geoespacial entre algunas fuentes de información, para así tener un consolidado de datos que funcionará como base en la intersección geoespacial con el evento ocurrido.

6.4.1 Modelo de datos relacional

A partir de llaves en cada conjunto de datos se realizó la unión y conglomeración de la información, el modelo de datos relacional en la **¡Error! No se encuentra el origen de la referencia.** se logra apreciar que en esta consolidación participaron los microdatos del DANE, los datos de las manzanas para obtener las coordenadas, y también .información sobre el IPM de las manzanas, en la Tabla 1 se puede identificar la fuente de cada uno de los datos que hacen parte del modelo relacional.

Tabla 1. Fuente de los datos del modelo relacional

Datos	Fuente
Vivienda	Censo Poblacional 2018 -DANE
Hogares	Censo Poblacional 2018 -DANE
Personas	Censo Poblacional 2018 -DANE
Georreferenciación	Censo Poblacional 2018 -DANE
Vulnerabilidad	Visor vulnerabilidad - DANE

6.4.2 Modelo de datos geoespacial:

Los datos que intervienen a nivel geoespacial están divididos en dos momentos del modelo, el primer momento hace referencia a la etapa de consolidación de la información, en donde se obtienen todos los datos bajo una misma tabla o base de datos unificada; el segundo momento es cuando el modelo está en producción, y es cuando ocurre el evento que el usuario ingresa el área afectada en forma de polígono, interactuando así con los polígonos de manzana y posteriormente con la

base de datos totalmente unificada, en la sección 6.6 se hace un mayor detalle del proceso.

En el anexo 2 se puede apreciar los datos que hacen parte del modelo geoespacial, las fuentes de estos datos provienen del sistema de datos abiertos de la alcaldía de Medellín.

6.5 Explotación de los datos y construcción del modelo

En esta etapa se buscó realizar un proceso de limpieza y consolidación de los datos para finalmente crear el modelo de consulta a los casos de estudio. En esta fase igualmente se definen las variables más relevantes según la literatura según el tipo de evento, Ortiz (2020) recalca que la integración de datos socio demográficos y ambientales son de gran importancia en la etapa de preparación, dado que permiten evaluar la vulnerabilidad de la población afectada. Algunos de los factores que se tienen en común indiferente al tipo de evento son el género, y edad de la población que permite establecer el dimensionamiento de servicios como asistencia sanitaria, vivienda y alimentación, como también la ubicación del área afectada para priorizar zonas para la entrega de las ayudas (Scarpin & Silva, 2014).

Las primeras etapas aquí descritas se pueden evidenciar en un repositorio [GitHub](#) para futuras actualizaciones dado que los datos del censo son del año 2018 y las dinámicas poblacionales pueden cambiar con el tiempo.

6.5.1 Lectura de los datos

Para la lectura de los datos, primero se requirió descargarlos de cada una de las fuentes de información debido a que estos datos se encontraban en datos abiertos nacional y municipal. Durante este proceso de ingesta se hizo necesario validar que los campos que servirían como llave estuvieran definidos bajo el mismo tipo de datos e igualmente bajo la misma composición para evitar dificultades en la consolidación posterior de los datos.

La mayoría de los datos vienen en formato csv, y otros en dbf, pero algunos tienen la disponibilidad de ser consumidos directamente desde la fuente de datos. No se eligió este método ya que los datos pueden ser eliminados de las fuentes o cambiar su estructura lo que afectaría posteriormente a la actualización de la información.

En la Tabla 2 se puede evidenciar cada uno de los datos que fueron cargados para la consolidación de información tanto a nivel relacional como geoespacial.

Tabla 2. Descripción de los datos cargados en la etapa de consolidación

Datos	Descripción	Formato
Viviendas	Información sobre las características de la vivienda a partir de los microdatos del censo.	Csv
Hogares	Información sobre las características de los hogares a partir del censo.	Csv
Personas	Información de cada una de las personas que respondieron al censo 2018	Csv

Georreferenciación	Información con el código de manzana que sirve para hacer unión con la información geoespacial	Csv
Vulnerabilidad	Datos que contienen el IPM	Dbf
Manzana	Descripción del mapa de manzanas del DANE que contiene las coordenadas de cada manzana.	Dbf
Violencia Intrafamiliar	Casos de violencia o seguridad a nivel geoespacial.	Csv
Llamada 123		Csv
Conducta contraria a la convivencia		Csv
Extorsión		Csv
Homicidio		Csv
Lesiones no fatales		Csv
Aprehensión		Csv
Secuestro		Csv
Hurto de personas		Csv
Hurto a residencias		Csv
Hurto Comercial		Csv

6.5.2 Criterios de limpieza de datos

Debido a que los datos obtenidos desde los sistemas de datos abiertos en ocasiones se encuentran sin ningún tratamiento de limpieza, se hizo necesario a partir de la exploración de los datos realizar unos criterios de limpieza para evitar datos atípicos o por fuera del área de cobertura del proyecto.

Se hizo necesario aplicar filtrado de la información del censo y del DANE en general para seleccionar el departamento de Antioquia y la ciudad de Medellín dado que los datos de esta fuente provienen a nivel departamental o nacional. Posteriormente cuando se hace una consolidación inicial de la información que será explicada en la sección 6.5.4.

En el caso de los datos de seguridad, se seleccionaron aquellos registros donde la ocurrencia del evento fue entre el 2017 y 2019 (Año hasta donde se cuenta con esta información) para tener un periodo de tiempo que abarque la cantidad de eventos ocurridos. Otro criterio de limpieza posterior fue eliminar los registros que no poseían coordenadas, o cuyas coordenadas eran inconsistentes con el rango de coordenadas de Medellín o incluso del país (En Colombia la longitud siempre es negativa).

6.5.3 Transformación de los datos

La información obtenida de los microdatos del censo 2018 se encuentra codificada a nivel de números, por lo que se hizo necesario transformar dicha codificación a los valores que representa cada uno. Para esto se utilizó el catálogo expuesto en la misma fuente de información en donde se exponía cada valor en cada pregunta que

representaba, y así para las columnas seleccionadas inicialmente se hizo el cambio para aquellas que aplicaban dado que algunos valores representaban cantidades. Los valores de cada una de la codificación se pueden evidenciar en el código R localizado en el repositorio [GitHub](#) en donde se presenta qué cambios se hicieron sobre la información obtenida.

Por otro lado, los datos de seguridad se transformaron a contadores para poder asociarlo a cada registro de manzana posteriormente en la consolidación geoespacial.

6.5.4 Consolidación de los datos por modelo relacional

Una vez se realizó la ingesta de los datos y hacer una selección inicial expuesta en la sección 6.5.2 se procedió con realizar los siguientes pasos a partir del modelo relacional:

- Primero se realizó un cruce entre los datos de manzana y con hogares a partir del código de la encuesta (COD_ENCUESTAS) ya que a partir de este era posible identificar las características del hogar de cada individuo.
- Posteriormente se realizó un nuevo cruce entre los datos obtenidos del paso anterior con la información de las viviendas a partir de la misma variable de código de la encuesta, para así obtener los datos de la vivienda y del hogar para cada individuo expuesto.
- Se realizó una selección de columnas de la base de georreferenciación para solo tener los datos agrupados por el código de la manzana y el código de la

encuesta para así de la consolidación anterior hacer una intersección con los campos seleccionados en este paso y poder tener para cada encuesta o individuo el código de la manzana a la cual pertenece.

- Finalmente, con el resultado anterior se hicieron dos cruces, una con el resultado anterior con la base de datos de las manzanas para poder obtener las coordenadas de cada manzana a partir del código de estas, y posteriormente con la base de datos obtenida en la consolidación geoespacial descrita en la sección 6.5.5 la cual para cada manzana se tiene el conteo de cada una de las variables de seguridad expuestas anteriormente.

6.5.5 Consolidación de los datos por modelo geoespacial

Para la consolidación a nivel geoespacial se usaron las coordenadas de cada uno de los eventos de seguridad. Inicialmente se definieron unas variables para acumular el contador para cada uno de los tipos de seguridad cargados. Para esto se calculó la distancia euclidiana entre las coordenadas de la manzana i , con las coordenadas del evento j y dado que el valor de la distancia resultante está dada a escala geoespacial se aplicó un factor de conversión para llevarlo a metros.

$$D_{i,j} = \frac{\sqrt{(long_i - long_j)^2 + (lat_i - lat_j)^2}}{0.000009009}$$

Una vez obtenida la distancia se procedía a evaluar si la distancia era menor a 100 metros para así solo obtener los eventos ocurridos en un radio de 100 metros del centroide de la manzana, si se cumplía con esta condición entonces hacia parte del

contador del evento evaluado, de esta manera para cada manzana se tenía la cantidad de eventos para cada uno de los temas de seguridad inicialmente cargados.

6.5.6 Selección de variables finales

Una vez obtenida la base de datos consolidada con las diferentes fuentes de información, se requirió de una etapa de selección de las variables a utilizar en el modelo de consulta cuando un evento ha ocurrido. A partir de la literatura se estableció todas las variables a utilizar, pero que según el tipo de evento el modelo seleccionará las variables que se visualizaran en los resultados generados.

6.5.7 Modelo de consulta y generación de resultados

El modelo de consulta definido para la generación de resultados es una función que requiere como dato de entrada el polígono del área afectada o que desea analizar en formato shape (shp), indicar el tipo de evento para la selección de las variables, y un área de influencia de manera opcional si el usuario desea abarcar un poco más a la población.

Una vez ingresado los requerimientos, el modelo realiza una lectura de la base de datos consolidada en los pasos anteriores (Esta base cabe aclarar que para esta base todos los pasos anteriores solo se repiten en caso de actualización), hace una lectura del archivo con el polígono, y si se asigna un área de influencia el modelo lo genera a partir del polígono ingresado.

Una vez realizado lo anterior, se procede con realizar una intersección geoespacial entre el polígono del área afectada y los polígonos de las manzanas del DANE; con esta intersección las manzanas se interceptan con la base de datos consolidada para así obtener solo los registros que aplican, y posteriormente selecciona las variables que se requieren analizar para el tipo de evento. Si el tipo de evento requiere tener conocimiento de zonas de espacio público que pueden ser estratégicos para ubicar la infraestructura necesaria en la atención del evento, el modelo realiza un análisis de distancias geográficas para identificar las zonas cercanas y exponerlo en los resultados.

Finalmente, con el resultado obtenido este es cargado en [GitHub](#) para que la interfaz de visualización definida pueda representar los datos obtenidos, y así finalmente el usuario o involucrados puedan tener todo el conocimiento posible de la población afectada y su entorno.

6.6 Arquitectura de datos en producción

En la Figura 1 se puede apreciar el paso a paso del funcionamiento del modelo propuesto. Los números en las flechas indican la secuencia del proceso de la arquitectura del modelo en producción.

1. Inicialmente cuando el usuario ingresa el área afectada, este genera una activación de la función de consulta y generación de resultados.
2. Al ser activada la función, esta realiza una ingesta de los datos suministrados y de los datos consolidados que se encuentran almacenados en [GitHub](#)

3. Con todos los datos obtenidos, el modelo se empieza a ejecutar en donde realiza una interacción del área afectada o ingresada con los datos consolidados y aplicando las condiciones internas del modelo.
4. Finalmente, al finalizar la ejecución del modelo este retorna unos resultados que son publicados para que la interfaz de visualización se actualice y así presente los resultados obtenidos para empezar a tomar decisiones de manera oportuna y ágil.

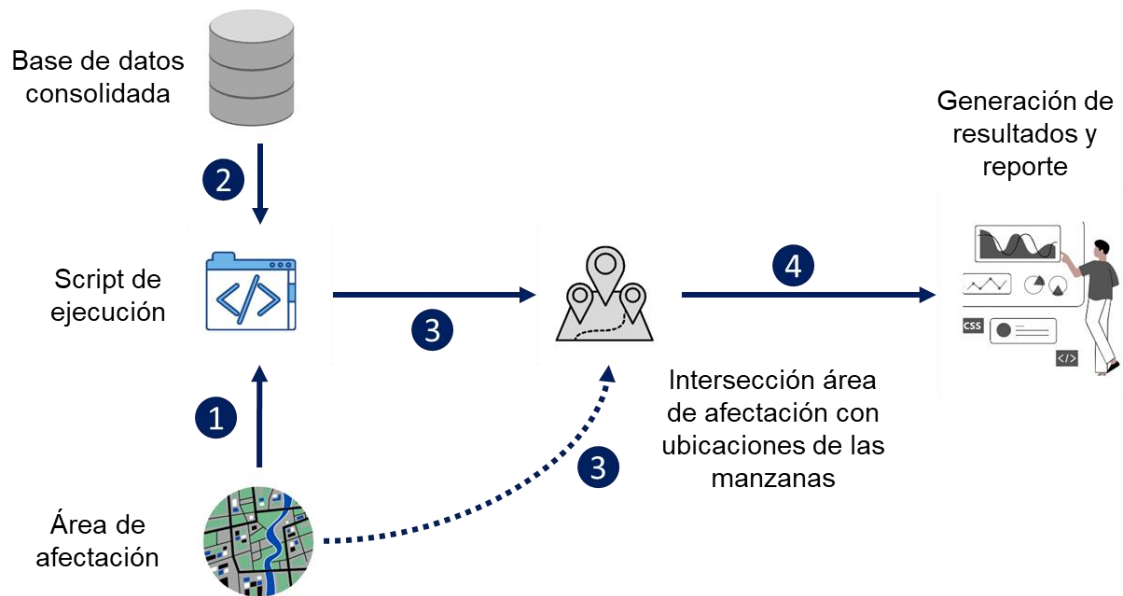


Figura 1. Arquitectura de datos en producción

6.7 Visualización de los resultados

Si bien la visualización de los resultados se puede implementar en cualquier herramienta de inteligencia de negocios como Tableau, Power-Bi, entre otros, o incluso desde un script diseñado para visualizar los resultados, en el presente trabajo se desarrolló el esquema de visualización en Power-Bi, para resumir los

resultados obtenidos en cada uno de los casos de estudio propuestos. El producto final se encuentra público en [GitHub](#) y adicionalmente se cuenta con el archivo de diseño original que puede ser modificado para quienes deseen hacer ajustes o nuevos aportes sobre lo ya construido.

Dado que la información que se debe suministrar para obtener los resultados, se propone cargar los datos y ejecutar el modelo en el [script](#) de R para así poder tener el resultado del modelo propuesto.

7 Casos de estudio

7.1 Planteamiento de Hipótesis

Existen dos hipótesis como medio de validación del modelo:

- La población seleccionada por el modelo es representativa del área de influencia afectada
- Existen diferencias en las necesidades de las personas según sus condiciones socioeconómicas ante la ocurrencia de un evento

Estas hipótesis son importantes para tener mayor soporte en las estimaciones de las características de las poblaciones y su entorno en las secciones 7.2, 7.3 y 7.4.

Para validar la primera hipótesis, se seleccionó aleatoriamente una de las manzanas afectadas por el evento y se utilizó uno de los fundamentos de la metodología KNN el cual es seleccionar las 5 manzanas más cercanas a la manzana de control.

Una vez realizado el proceso de selección de la manzana afectada y sus 5 manzanas cercanas se procedió a realizar segmentaciones por algunas de las variables de la población y su entorno, entre estas variables fueron el rango de edad, el género, el estrato, la fuente de ingreso, fuente de agua para preparación de alimentos, el hurto a personas y eventos de violencia familiar. Para cada una de estas segmentaciones se realizaron pruebas chi² en donde se buscaba validar que las proporciones de cada segmentación eran similares entre la manzana de control y sus cercanas.

Tabla 3. Valores P prueba Chi²

Variable	Valor-P
Rango de edad	0.77
Genero	0.66
Estrato	6.17e-07
Fuente de ingreso	0.28
Agua para los alimentos	0.82
Hurto a Personas	0.99
Violencia Familiar	< 2.2e-16

En la Tabla 3 se aprecian los resultados obtenidos en las pruebas estadísticas de las proporciones, de estos se logra identificar que variables como el estrato y la violencia familiar pueden presentar diferencias entre manzanas, esto se debe a que existen zonas donde el estrato socioeconómico de una zona puede ser variado, o

incluso de una manzana a otra pasar de estrato 2 a 3 por ejemplo. En el tema de la violencia familiar las diferencias se pueden dar porque son comportamientos más individualizados que generalizados como los eventos de hurto a personas cuya área de influencia puede ser mayor a la de la violencia familiar.

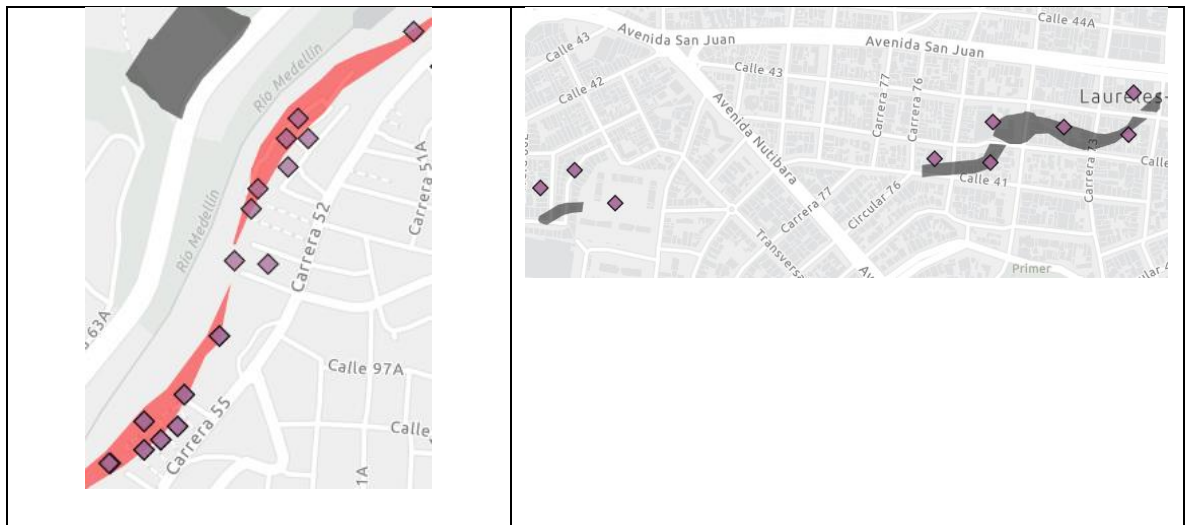
Para la segunda hipótesis planteada, estas se pueden validar a partir de los análisis de casos de estudio en las secciones posteriores, en donde se logra evidenciar diferencias a nivel de la población, sus viviendas y entorno cuando se comparan entre una zona con bajo nivel de ingresos contra una zona de mayor nivel de ingreso.

7.2 Caso Inundación

Para este caso de estudio se planteó un escenario de inundación en dos zonas donde sus condiciones de vida son totalmente diferentes, para poder seleccionar estas zonas se tuvo en cuenta las encuestas de calidad de vida, en donde se analizaron a nivel de barrios el promedio de los ingresos de las viviendas, el estrato socioeconómico predominante, el acceso a servicios públicos, condiciones de la vivienda y el acceso a los alimentos diarios; y que a su vez estuvieran expuestos a zonas de riesgo de inundación para analizar el escenario.

Tabla 4. Zonas seleccionadas para el caso de estudio

Zona 1	Zona 2
--------	--------



En la Tabla 4 se puede apreciar las zonas seleccionadas, la zona 1 se encuentra localizada cerca al Río Medellín en su margen oriental a la altura de la estación tricentenario del metro, y la Zona 2 se localizan en el barrio laureles.

Una vez ejecutado el modelo bajo el escenario de inundación propuesto se obtuvieron los siguientes resultados:

Frente al componente de características de la población se logra observar que la cantidad de personas afectadas tanto para la zona 1 como para la zona 2 son cercanas (2.043 y 2.358 respectivamente), pero en la zona 1 es mayor el número de hombres con respecto a las mujeres (1.039 y 1.004), y en la zona 2 esta participación cambia (1.023 hombres y 1.335 mujeres). En términos de rangos etarios la zona 1, el 25% de la población está entre los 20 y 29 años y 27% son menores de 19 años, mientras que en la zona 2, el 65% de los habitantes esta entre los 20 y 70 años, y 14 % son menores de 19 años, presentándose entonces una menor cantidad de niños en dicha zona que también se puede corroborar cuando

se mira la participación de los parentescos en donde la zona 1 los hijos son los de mayor participación y en la zona 2 son los jefes de hogar, esto también puede influir en el resultado de personas que poseen dificultades de salud en su vida diaria ya que la zona 2 posee una población más envejecida.

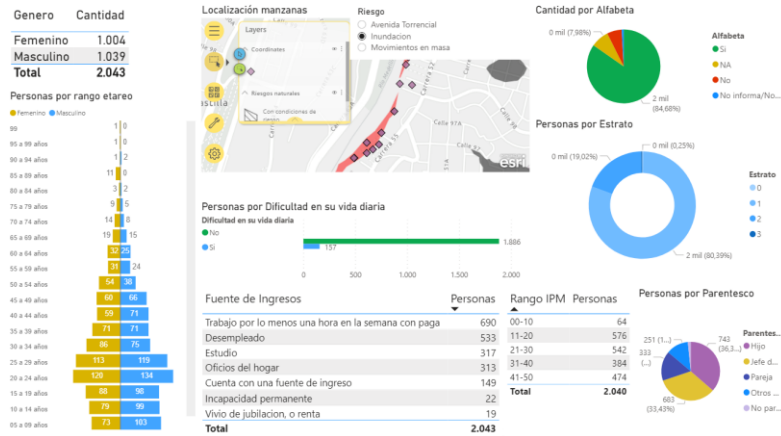


Figura 2. características de las personas de la zona 1

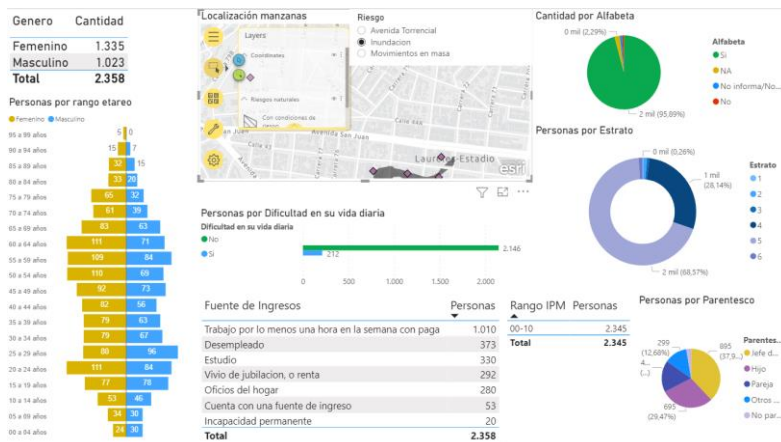


Figura 3. características de las personas de la zona 2

A nivel de Índice de Pobre Multidimensional (IPM), mientras en la zona 1 existen manzanas con varios niveles, incluso por encima de 30, la zona 2 solo se localiza

en el primer nivel el cual es un indicador positivo, lo que se logra identificar las diferencias en niveles de pobreza de las dos zonas. A nivel de educación o alfabetismo mientras en la zona 2 el 96% de las personas son alfabetas independiente de su nivel de educación en la zona 1 este porcentaje disminuye a un 85% lo que podría enmarcar una diferencia en los accesos y oportunidades a la educación y al trabajo. A nivel del estrato de la vivienda que a su vez puede indicar una relación directa con los ingresos mensuales de las personas, el estrato predominante en la zona 1 y 2 son el estrato 1 y 5 respectivamente, aquí se puede apreciar que los ingresos de las personas de la zona 2 son superiores a los de la zona 1 lo que podría influir en su calidad de vida y que se puede corroborar con su fuente de ingresos ya que en la zona 2 el 28% de la población es desempleada y el 14% está dedicada a sus estudios comparada con la población de la zona 1 donde el 42% de las personas son desempleadas y sin fuentes de ingreso, y el 15% está en etapa de estudios dejando solo un porcentaje de personas con fuente de ingresos o empleo inferior al porcentaje de la zona 2.

Pasando al frente de las características de las viviendas de las personas para las dos zonas, aunque en ambas zonas existen personas en el hogar entre los rangos 1-4, 5-8 y 9-12, para la zona 1 es mayor el porcentaje de la segunda categoría e incluso aparece la categoría de 9-12 personas en el hogar esto indica que en la zona 1 se puede presentar al interior de las vivienda hacinamiento de personas. Frente al tipo de vivienda en la zona 1 es mayor el porcentaje de personas que viven en casa o en cuarto lo que podría ser un indicador en temas de inundación que

existan personas más expuestas a daños materiales por inundación en los niveles bajos, ya que para apartamentos es posible que estos sean edificaciones de más de 2 pisos y por ende menos exposición a daños por inundación. En temas de acceso a servicios públicos mientras en la zona 2 casi el 100% de las personas cuenta con los servicios públicos básicos, en la zona 1 existen personas sin acceso a acueducto, alcantarillado y recolección de basuras por lo que estos individuos podrían estar más expuestos a temas de salud y dicha exposición aumentarse en momentos de inundación que puedan generar problemas de salubridad.

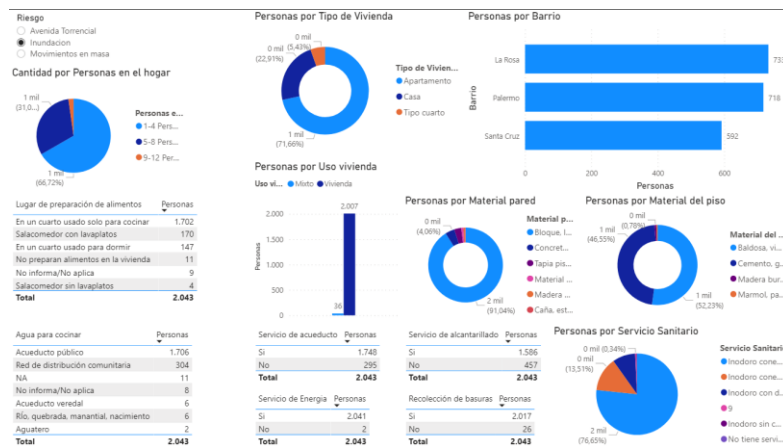


Figura 4. características vivienda de la población de la zona 1

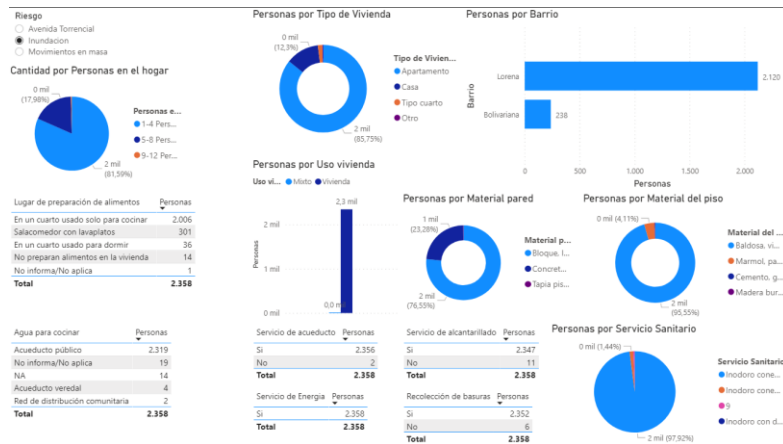


Figura 5 características vivienda de la población de la zona 2

Al analizar el tema de preparación de alimentos que es un tema importante en la calidad para la salud de las personas la zona 2 casi el 100% de las personas lo realizan en zonas adecuadas para su función y con acueducto público, mientras que en la zona 1 las opciones son diversas como por ejemplo realizar la preparación de alimentos en zonas que son para dormir y con un acceso irregular al acueducto para la preparación de los mismos, esto influye en la calidad de los alimentos por la higiene que estos puedan presentar. Finalmente llama la atención el material de las viviendas, ya que en la zona 2 tanto el piso como las paredes están construidos en materiales de mejor calidad en términos estéticos e incluso a nivel de estabilidad de la estructura, en la zona 1 los materiales de las paredes varían y en el piso su material predominante es el cemento, por lo que a nivel estético sus viviendas podrían estar en “obra gris” o no tener la obra blanca totalmente terminada. Estos hallazgos en este frente pueden evidenciar que ante un evento de inundación en términos de vivienda las personas de la zona 1 requieren un mayor acompañamiento en la recuperación y adecuación de sus viviendas que permitan

mejorar su calidad de vida e incluso según las etapas de atención de desastres buscar el objetivo de que en la etapa de recuperación esta busque llegar a su nivel anterior o mejor en temas de calidad. Mientras que para la zona 2 el nivel de intervención y acompañamiento puede ser menor dado que estos habitan en mejores condiciones.

Finalmente, en temas del entorno relacionados con seguridad, al traer la cantidad de eventos ocurridos en tres años se observa que mientras en la zona 1 el hurto a personas está por debajo de 40 eventos, en la zona 2 son más de 40 ocurrencias, lo que podría darse dado que en la zona 2 los niveles de ingresos y calidad de vida son mejores por lo que se convierten en objetivos más apetecibles para este tipo de criminalidad.

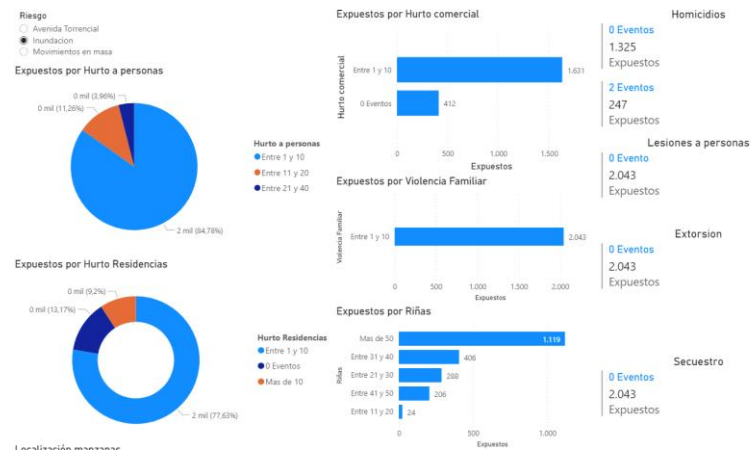


Figura 6. características del entorno de la zona 1

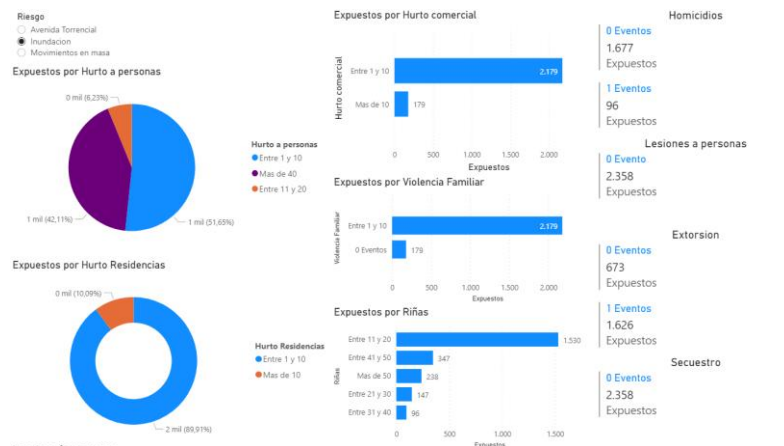


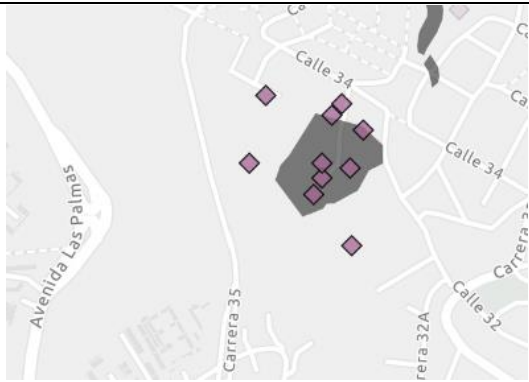
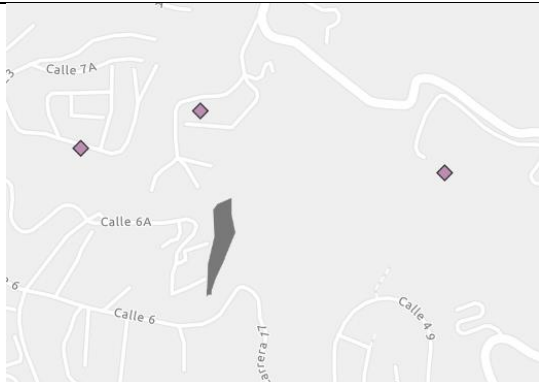
Figura 7. características del entorno de la zona 2

De igual forma el nivel de hurto a residencias pareciese más frecuente en la zona 2 que en la zona 1 que presenta participación de cero eventos para cada individuo, pero al mirar el hurto comercial este es más predominante en la zona 1 que en la 2, como también el número de expuestos a riñas, donde hay exposición a más de 50 riñas en 3 años este tipo de violencia puede generar vulnerabilidad de las personas en tiempos de recuperación en donde se puedan generar riñas o encuentros incómodos entre los habitantes, ya que a nivel de violencia familiar ambas zonas presentan la misma proporción de expuestos, pero en temas de extorsión la zona 2 presenta algunos expuestos a estos eventos. Estos hallazgos son importantes para conocer el nivel de vulnerabilidad social que las personas presentaban antes de la ocurrencia del evento y como dentro de las estrategias de recuperación y acompañamiento se les pueda agregar primero un valor agregado en temas de fortalecimiento social, y también en las consideraciones de seguridad que se deben tener en la logística de entrega de ayudas a las personas, o si las ayudas deben también mirar hacia otros objetivos sociales.

7.3 Caso Deslizamiento

Para este caso de estudio se planteó un escenario de deslizamiento o movimiento en masa inicialmente para dos zonas donde sus condiciones de vida son totalmente diferentes tal como se seleccionaron para el caso de inundación, pero al observar esta exposición se halló que los barrios de estratos bajos son los que más niveles de exposición tienen a los riesgos de evento natural.

Tabla 5. Zonas seleccionada caso deslizamiento

Zona 3	Zona 4
	

En la Tabla 5 se puede apreciar las zonas seleccionadas, la zona 3 se encuentra localizada en el barrio la asomadera, al oriente de la avenida las palmas, y la Zona 4 se localiza en el barrio las palmas y el tesoro.

Una vez ejecutado el modelo bajo el escenario de deslizamiento propuesto se obtuvieron los siguientes resultados:

Frente al componente de características de la población se logra observar que la cantidad de personas afectadas tanto para la zona 3 como para la zona 4 difieren ya que en la zona 3 son 2.741 personas expuestas y en la zona 4 son 1.114. En

términos de género en la zona 3 es mayor el número de hombres con respecto a las mujeres (1.375 y 1.366 respectivamente), y en la zona 4 esta participación cambia (503 hombres y 611 mujeres). Observando la distribución por rangos etarios en la zona 3, el 25% de los habitantes esta entre los 20 y 29 años, y el 30% son menores de 19 años, mientras que en la zona 4 el 44 % de los habitantes esta entre los 25 y los 54 años, y el 27% son menores de 19 años, lo que podría indicar que es una zona de crecimiento familiar.

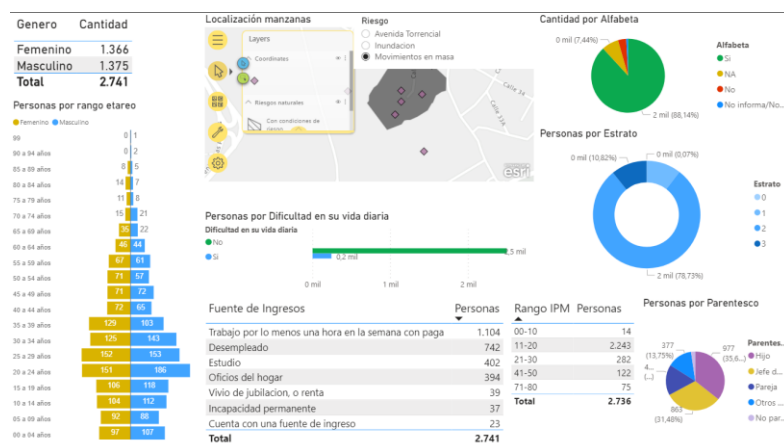


Figura 8. Características de la población de la zona 3

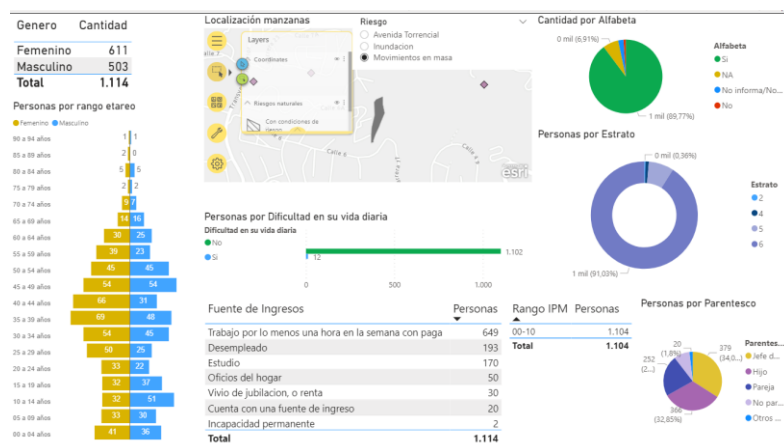


Figura 9. Características de la población de la zona 4

A nivel de Índice de Pobre Multidimensional (IPM), mientras en la zona 3 existen manzanas con varios niveles, incluso en rangos superiores a 40, la zona 4 solo se localiza en el primer nivel el cual es un indicador positivo, lo que se logra identificar las diferencias en niveles de pobreza de las dos zonas, en donde la zona 3 tiene personas que viven en altos niveles de pobreza. A nivel del estrato de la vivienda que a su vez puede indicar una relación directa con los ingresos mensuales de las personas, el estrato predominante en la zona 3 y 4 son el estrato 1 y 6 respectivamente, aquí se puede apreciar que los ingresos de las personas de la zona 4 son superiores a los de la zona 3 lo que podría influir en su calidad de vida y que se puede corroborar con su fuente de ingresos ya que en la zona 4 el 22% de la población es desempleada y el 15% está dedicada a sus estudios comparada con la población de la zona 3 donde el 43% de las personas son desempleadas y sin fuentes de ingreso, y el 15% está en etapa de estudios dejando solo un porcentaje de personas con fuente de ingresos o empleo inferior al porcentaje de la zona 4.

Frente a las características de las viviendas de las personas para las dos zonas, aunque en ambas zonas existen personas en el hogar entre los rangos 1-4, 5-8, para la zona 3 es mayor el porcentaje de la segunda categoría e incluso aparece la categoría de 9-12 personas en el hogar esto indica que en la zona 3 se puede presentar al interior de las vivienda hacinamiento de personas. Frente al tipo de vivienda en la zona 3 es mayor el porcentaje de personas que viven en casa o en cuarto lo que podría ser un indicador en temas de deslizamiento que existan personas más expuestas a daños materiales en los niveles bajos, o incluso en la

pérdida de vidas humanas, ya que para apartamentos es posible que estos sean edificaciones de más de 2 pisos y por ende menos exposición a daños. En temas de acceso a servicios públicos mientras en la zona 4 casi el 100% de las personas cuenta con los servicios públicos básicos, en la zona 3 existen personas sin acceso a acueducto, alcantarillado y recolección de basuras por lo que estos individuos podrían estar más expuestos a temas de salud.



Figura 10. características de las viviendas de la zona 3

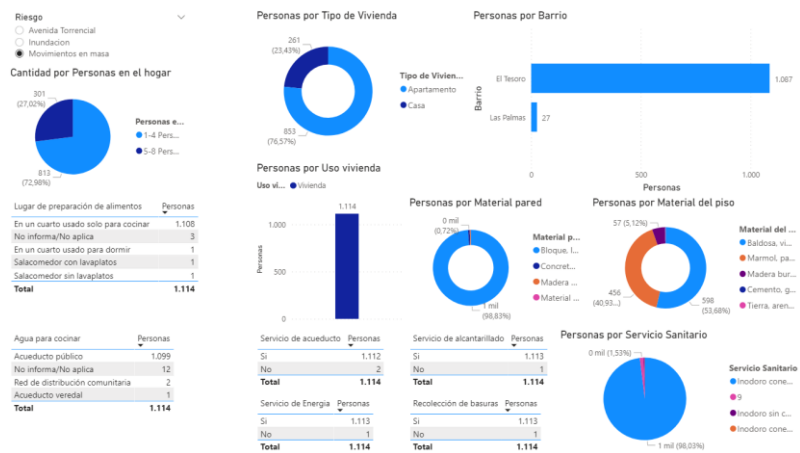


Figura 11. características de las viviendas de la zona 4

Al analizar el tema de preparación de alimentos que es un tema importante en la calidad para la salud de las personas la zona 4 casi el 100% de las personas lo

realizan en zonas adecuadas para su función y con acueducto público, en la zona 3 las opciones son diversas como por ejemplo realizar la preparación de alimentos en zonas que son para dormir y con un acceso irregular al acueducto para la preparación de los mismos, esto influye en la calidad de los alimentos por la higiene que estos puedan presentar. Finalmente llama la atención el material de las viviendas, ya que en la zona 4 tanto el piso como las paredes están contruidos en materiales con más “lujo” a nivel estético e incluso a nivel de estabilidad de la estructura, en la zona 3 los materiales de las paredes varían y en el piso su material predominante es la baldosa, aunque también hay en cemento, por lo que a nivel estético sus viviendas podrían estar en “obra gris” o no tener la obra blanca totalmente terminada. Estos hallazgos en este frente pueden evidenciar que ante un evento de deslizamiento en términos de vivienda las personas de la zona 3 requieren un mayor acompañamiento en la recuperación y adecuación de sus viviendas que permitan mejorar su calidad de vida e incluso según las etapas de atención de desastres buscar el objetivo de que en la etapa de recuperación esta busque llegar a su nivel anterior o mejor en temas de calidad. Mientras que para la zona 4 el nivel de intervención y acompañamiento puede ser menor dado que estos habitan en mejores condiciones.

Finalmente, en temas del entorno relacionados con seguridad, al traer la cantidad de eventos ocurridos en tres años se observa que a pesar de que ambos están expuestos a hurtos a personas, en la zona 3 predomina la categoría “Entre 1 y 10” mientras que en la zona 4 predomina la categoría “0 Eventos”, este hallazgo es

contrario a lo observado en el caso de inundación, y puede estar influenciado al tránsito de las personas en las calles, donde posiblemente en la zona 4 este tránsito puede ser inferior al de la zona 2 en el caso de inundación.

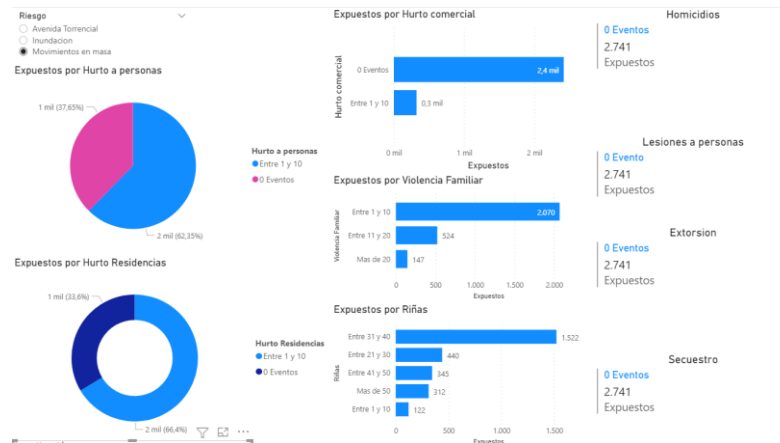


Figura 12. características del entorno de la población de la zona 3

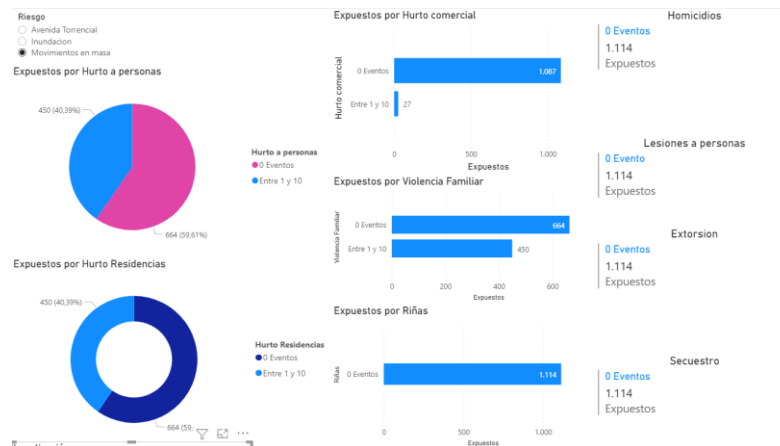


Figura 13. características del entorno de la población de la zona 4

De igual forma el nivel de hurto a residencias es más frecuente la ocurrencia de estos eventos en la zona 3 que en la 4, quizás se deba a los controles de seguridad que se manejan en los ingresos a los apartamentos o unidades residenciales de la zona 4. En temas de riñas, en la zona 4 no se tienen registros de este tipo de

incidencia en un periodo de tres años, mientras que la zona 3 ha llegado a tener hasta 40 eventos de riñas en sus alrededores lo que evidencia un tema de vulnerabilidad social en dicha zona. Por último, llama la atención el tema de violencia familiar, si bien ambas zonas han estado expuestas a este tipo de eventos en sus alrededores, la zona 3 presenta categorías superiores lo que termina de corroborar ese nivel de vulnerabilidad social que pueden presentar las personas en esta.

7.4 Caso avenida torrencial:

Para este caso de estudio se planteó un escenario de avenida torrencial en dos zonas donde sus condiciones de vida son totalmente diferentes, para poder seleccionar estas zonas se usó las encuestas de calidad de vida las cuales fueron agrupadas y caracterizadas a nivel de barrio. De igual forma se consideró el estrato socioeconómico predominante de la zona y así proceder con el modelo y su análisis correspondiente. Las Avenidas torrenciales son eventos que ocurren en los cauces de los afluentes hídricos, en donde traen sedimentos y escombros que pueden generar daños a las viviendas y poner en riesgo la vida de los afectados.

Tabla 6. Zonas seleccionada caso avenida torrencial

Zona 5	Zona 6
--------	--------

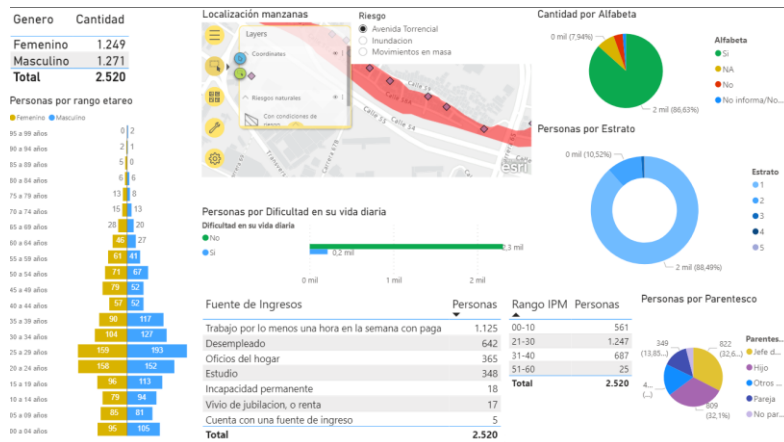


Figura 14. Características de la población de la zona 5

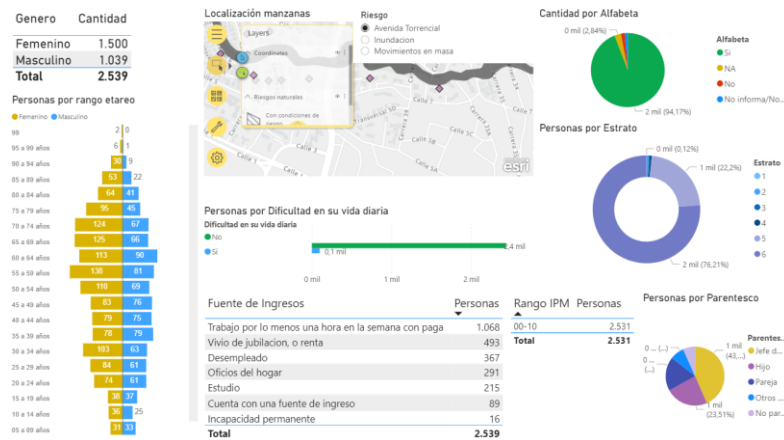


Figura 15. Características de la población de la zona 6

A nivel de Índice de Pobre Multidimensional (IPM), mientras en la zona 5 existen manzanas con varios niveles, incluso en rangos superiores a 40, la zona 6 solo se localiza en el primer nivel el cual es un indicador positivo, lo que se logra identificar las diferencias en niveles de pobreza de las dos zonas, en donde la zona 5 tiene personas que viven en altos niveles de pobreza. A nivel del estrato de la vivienda que a su vez puede indicar una relación directa con los ingresos mensuales de las personas, el estrato predominante en la zona 5 y 6 son el estrato 1 y 6 respectivamente, aquí se puede apreciar que los ingresos de las personas de la

zona 6 son superiores a los de la zona 5 lo que podría influir en su calidad de vida y que se puede corroborar con su fuente de ingresos ya que en la zona 6 el 26% de la población es desempleada y el 8% está dedicada a sus estudios comparada con la población de la zona 5 donde el 41% de las personas son desempleadas y sin fuentes de ingreso, y el 14% está en etapa de estudios dejando solo un porcentaje de personas con fuente de ingresos o empleo inferior al porcentaje de la zona 6.

Frente a las características de las viviendas de las personas para las dos zonas, aunque en ambas zonas existen personas en el hogar entre los rangos 1-4, 5-8, para la zona 5 es mayor el porcentaje de la segunda categoría e incluso aparece la categoría de 9-12 personas y más de 12 personas en el hogar, esto indica que en la zona 5 se puede presentar al interior de las vivienda hacinamiento de personas. Frente al tipo de vivienda en la zona 5 es mayor el porcentaje de personas que viven en casa o en cuarto con respecto a la zona 6, lo que podría ser un indicador en temas de avenida torrencial que existan personas más expuestas a daños materiales en los niveles bajos, o incluso en la pérdida de vidas humanas, ya que para apartamentos es posible que estos sean edificaciones de más de 2 pisos y por ende menos exposición a daños. En temas de acceso a servicios públicos mientras en la zona 6 casi el 100% de las personas cuenta con los servicios públicos básicos, en la zona 5 existen personas sin acceso a acueducto, alcantarillado y recolección de basuras por lo que estos individuos podrían estar más expuestos a temas de salud y dicha exposición aumentarse en momentos de avenida torrencial que puedan generar problemas de salubridad.

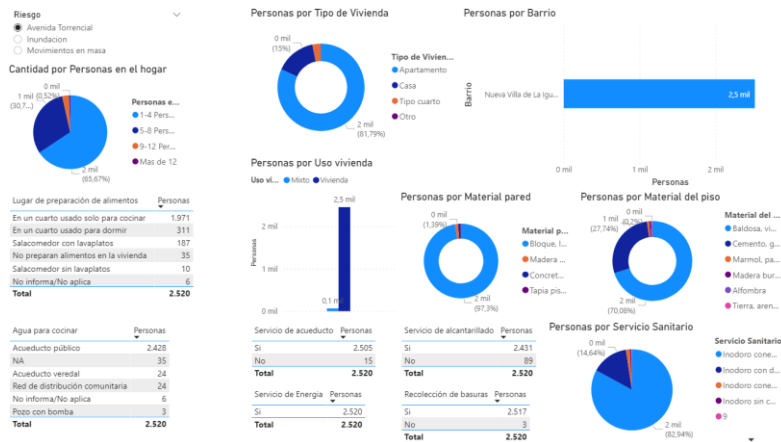


Figura 16. Características de la vivienda de la zona 5

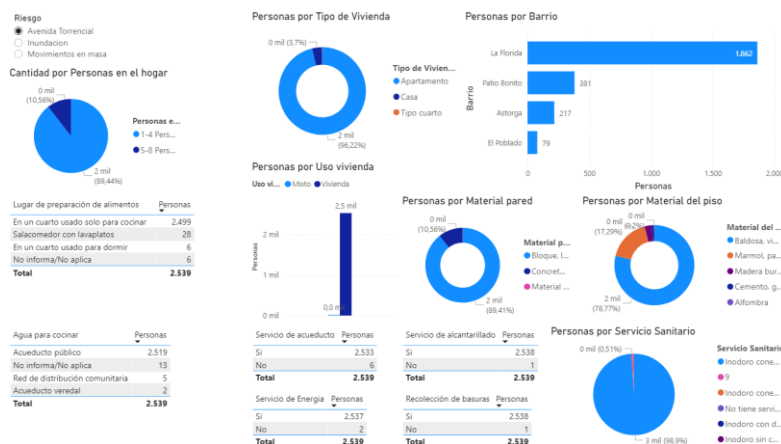


Figura 17. Características de la vivienda de la zona 6

Al analizar el tema de preparación de alimentos que es un tema importante en la calidad para la salud de las personas la zona 6 casi el 100% de las personas lo realizan en zonas adecuadas para su función y con acueducto público, en la zona 5 las opciones son diversas como por ejemplo realizar la preparación de alimentos en zonas que son para dormir y con un acceso irregular al acueducto para la preparación de los mismos, esto influye en la calidad de los alimentos por la higiene que estos puedan presentar, estos hallazgos también se encontraron en los casos de inundación y deslizamientos. Finalmente, el material de las viviendas, se observa

que para la zona 6 el piso como las paredes están contruidos en materiales con más “lujo” a nivel estético e incluso a nivel de estabilidad de la estructura, mientras que en la zona 5 los materiales de las paredes varían y en el piso su material predominante es la baldosa, aunque también hay en cemento, por lo que a nivel estético sus viviendas podrían estar en “obra gris” o no tener la obra blanca totalmente terminada. Estos hallazgos en este frente pueden evidenciar que ante un evento de avenida torrencial en términos de vivienda las personas de la zona 5 requieren un mayor acompañamiento en la recuperación y adecuación de sus viviendas que permitan mejorar su calidad de vida e incluso según las etapas de atención de desastres buscar el objetivo de que en la etapa de recuperación esta busque llegar a su nivel anterior o mejor en temas de calidad. Mientras que para la zona 6 el nivel de intervención y acompañamiento puede ser menor dado que estos habitan en mejores condiciones.

Finalmente, en temas del entorno relacionados con seguridad, al traer la cantidad de eventos ocurridos en tres años se observa que a pesar de que ambos están expuestos a hurtos a personas, en la zona 6 es mayor la proporción de expuestos a más de 40 eventos comparado con la zona 5, este punto se puede dar dado que la zona 6 es un área con mucha afluencia de personas, y por su alto nivel de ingresos llama la atención para la ocurrencia de este tipo de eventos.

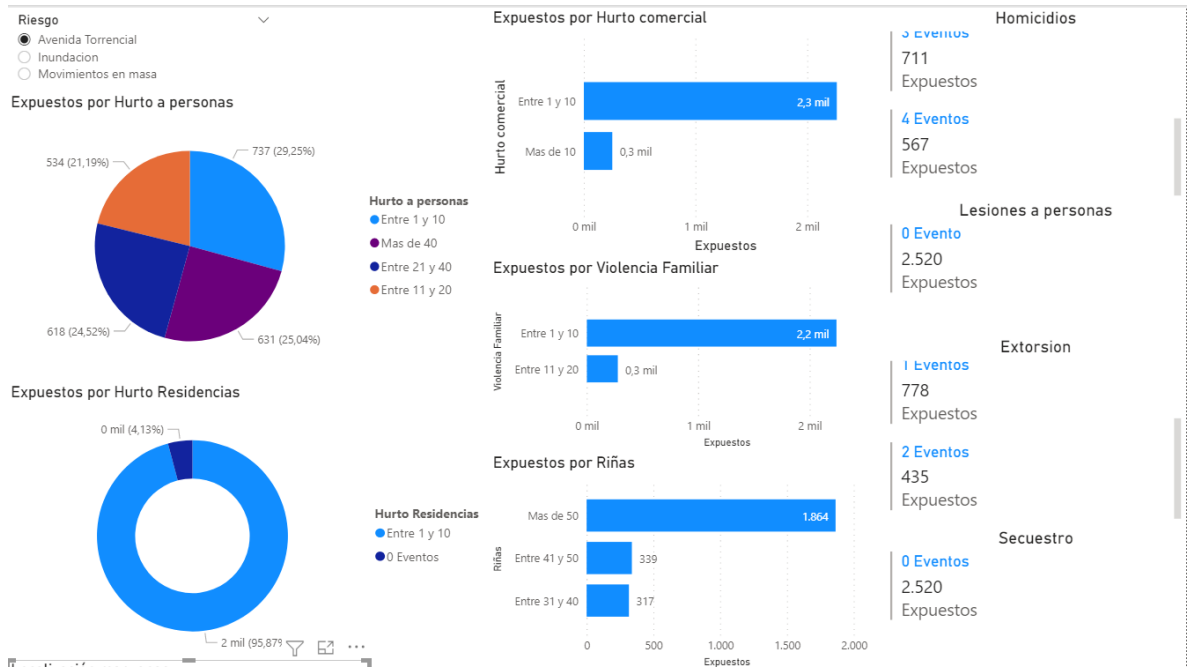


Figura 18. Características del entorno de la zona 5

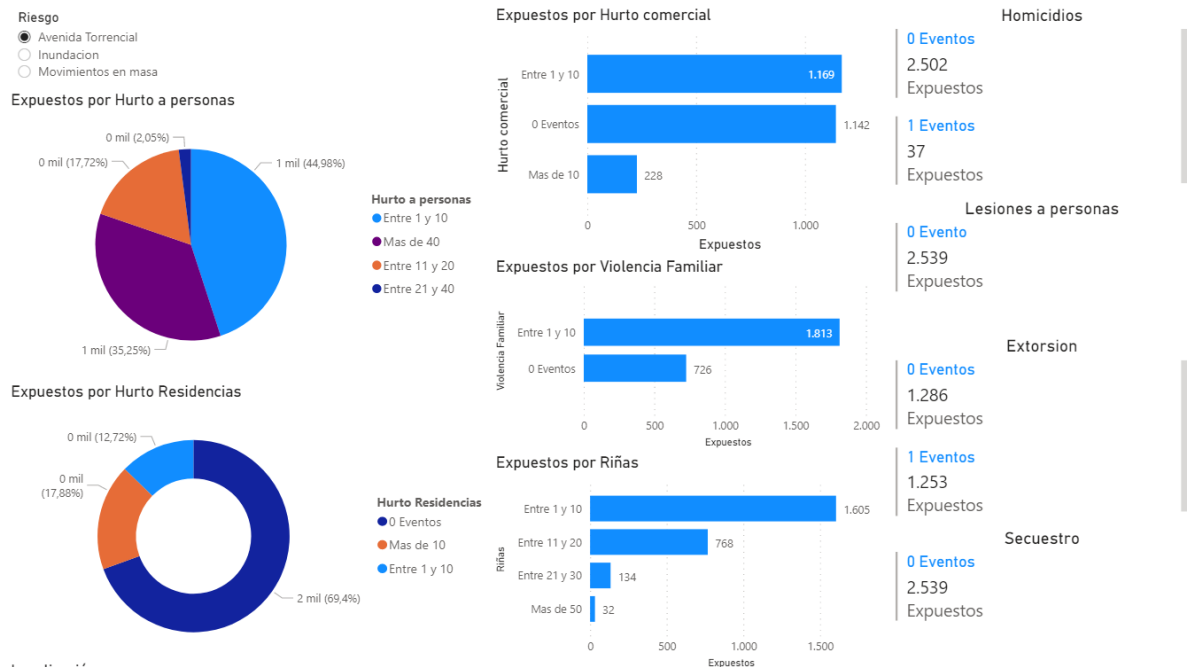


Figura 19. Características del entorno de la zona 6

Con respecto al nivel de hurto a residencias si bien en la zona 6 hay exposición a más de 10 eventos, predomina la categoría “0 eventos”, mientras que para la zona 5 predomina la categoría “Entre 1 y 10”. De igual forma como la zona 6 es una zona comercial y residencial esta presenta mayor número de expuestos a hurtos comerciales con respecto a la zona 5, en donde no hay tanto comercio a sus alrededores. Por otro lado, aunque ambas zonas están expuestas a riñas, la proporción en la categoría de más de 50 eventos en un periodo de tres años es superior en la zona 5, como también el tema de los homicidios, en donde han ocurrido más de estos eventos si se comparan con la zona 6.

8 Discusión y Conclusiones

Una primera discusión está basada en los resultados anteriores en donde es evidente primero las diferencias en la calidad de vida de las personas cuando se comparan por ubicación y estrato, estos hallazgos permiten entender que frente a eventos de desastres si bien ambas poblaciones requieren atención y acompañamiento el nivel de profundidad o estrategia de acompañamiento debe ser totalmente diferente dadas las necesidades que cada población presenta a nivel de su hábitat y su entorno, esto se puede corroborar con las características de su vivienda, los niveles de desempleo o las fuentes de ingresos en el hogar por lo que el acompañamiento y gestión a desastres debe ser más profundo en poblaciones con necesidades mayores no solo en la recuperación de sus bienes materiales, sino también en el fortalecimiento social de las personas y de su calidad de vida.

El entorno de las poblaciones es relevante, se pudo evidenciar como en algunas zonas el nivel de seguridad puede ser bajo comparado con otras zonas donde los eventos homicidios, extorsión eran bajos, pero que a su vez podían presentar un comportamiento contrario en temas de hurto a personas donde los barrios de mayor nivel de ingresos y afluencia de personas caminando eran condiciones que podrían estar influyendo en el alto número de eventos de este tipo de crimen. Estos hallazgos a nivel del entorno son importantes debido a que, si bien ambas poblaciones requieren atención ante la ocurrencia de un desastre, dentro de sus estrategias de acompañamiento debe considerarse un frente de mejoramiento en las condiciones sociales y de seguridad, de igual forma tener esta información permite tener estrategias en la protección de quienes van acudir la emergencia y a su vez en la seguridad de los insumos durante la fase de atención a las poblaciones.

Finalmente el modelo desarrollado tiene la capacidad de cuantificar y caracterizar a una población afectada en términos del individuo, su vivienda y su entorno, generando así información relevante en el diseño de un plan de logística humanitaria, ya que este aporta una aproximación a la etapa de censo de personas y viviendas afectadas, y a la implementación de estrategias de atención y acompañamiento de la población.

9 Trabajos futuros

Para trabajos futuros se recomienda robustecer la información social y del entorno, porque a la fecha si bien existe esta información a nivel de comuna mediante la encuesta de calidad de vida, es necesario llevar u obtener esta información a un

nivel de granularidad mayor. También se propone escalar el modelo a otras regiones, o ampliarlo a toda el área metropolitana del valle de aburra, ya que a medida que el territorio crece las fronteras entre municipios son cada vez menos claras, por lo que se hace necesario poder obtener una mayor área de estudio.

10 Aspectos éticos

Los datos abiertos utilizados en el presente trabajo, beneficiará las poblaciones que pueden estar implicadas o con mayor exposición a los escenarios aquí expuestos, como también al gobierno local en el apoyo en temas de logística humanitaria que son relevantes a la hora de tomar acciones rápidas y certeras. Se espera que este modelo sea utilizado por las partes interesadas, especialmente el gobierno local, o personas con un buen manejo de los datos que incluso pueda tomar estos datos como un nuevo punto de partida para otro modelo, o mejoras de este.

11 Uso de los datos

11.1 Políticas de uso

El modelo expuesto puede ser utilizado para escalar a otro ámbito o población de interés de estudio que se quiera llevar para ampliar la red de conocimiento y ayuda en momentos de crisis como lo son la atención de eventos de inundación y deslizamientos en una población.

11.2 Archivación de datos y modelo

Los datos consolidados y el modelo se encuentran disponibles en un repositorio GitHub para futuros desarrollos o análisis.

12 Bibliografía

- Abazari, S. R., Aghsami, A., & Rabbani, M. (2020). Prepositioning and distributing relief items in humanitarian logistics with uncertain parameters. *Socio-Economic Planning Sciences*.
- Abushaikha, I., & Schumann-Bölsche, D. (2016). Mobile phones: Established technologies for innovative humanitarian logistics concepts. *Procedia Engineering*, 191-198.
- Alcaldia de Medellin. (2013). *POT*. Medellín: Alcaldia de Medellin.
- Banco Mundial. (2012). *Análisis de la gestión del riesgo de desastres en Colombia: un aporte para la construcción de políticas públicas*.
- Banco Mundial. (15 de 09 de 2019). *Crecimiento de la Población*. Obtenido de https://datos.bancomundial.org/indicador/SP.POP.GROW?locations=MX&most_recent_value_desc=true
- Banco Mundial. (11 de 12 de 2020). *Banco Mundial*. Obtenido de <https://www.bancomundial.org/es/news/feature/2014/02/03/hackathons-tech-savvy-innovators-asia-disaster-resilience>
- Cinnamon, J. (2020). *Humanitarian Mapping*. Toronto: Elsevier.
- Crumbly, J., & Carter, L. (2015). Social media and humanitarian logistics: The impact of task technology fit on new service development. *Procedia Engineering*, 412-416.

- DANE. (01 de 01 de 2021). *Colombia- Censo Nacional de Población y Vivienda*.
Obtenido de http://microdatos.dane.gov.co/index.php/catalog/643/get_microdata
- Díaz-Delgado, C., & Iniestra, J. G. (2014). Flood Risk Assessment in Humanitarian Logistics Process Design. *Journal of Applied Research and Technology*, 976-984.
- Dinh, T. D., & Pham, V. M. (2020). Studying Urban Expansion and Landscape Surrounding Monuments for Conservation the World Cultural Heritage in Hue City—View from GIS and Remote Sensing. *AUC 2019*, 317-331.
- Galleguillos-Pozo, R., Jordan-H, E. P., & Garcia, M. V. (2019). Aplicación Lógica Difusa y AHP en caso de Estudio Ecuatoriano de Logística humanitaria. *Iberian Journal of Information Systems and Technologies*, 433–444.
- Gómez, D. M., Sarache, W., & Trujillo, M. (2017). Identificación y Análisis de una Red de Ayuda Humanitaria.Un Caso de Estudio. *Información Tecnológica*, 115-124.
- Gupta, S., Altay, N., & Luo, Z. (2019). Big data in humanitarian supply chain management: a review and further research directions. *Ann Oper Res*, 1153–1173.
- Jabbour, C. J., Sobreiro, V. A., Jabbour, A. B., Campos, L. M., Mariano, E. B., & Renwick, D. W. (2019). An analysis of the literature on humanitarian logistics

and supply chain management: paving the way for future studies. *Ann Oper Res*, 289-307.

Kanga, S., Meraj, G., Sudhanshu, M. F., & M. S. Nathawat, S. K. (2021). Analyzing the Risk to COVID-19 Infection using Remote Sensing and GIS. *Risk Analysis*, 801-813.

Khan, M., Lee, H. Y., & Bae, J. H. (2019). The Role of Transparency in Humanitarian Logistics. *Sustainability*.

Kilicoglu, C., Cetin, M., Aricak, B., & Sevik, H. (2021). Integrating multicriteria decision-making analysis for a GIS-based settlement area in the district of Atakum, Samsun, Turkey. *Theoretical and Applied Climatology volume* , pages379–388.

Ochoa, A., Rudomin, I., Vargas-Solar, G., Espinosa-Oviedo, J. A., Pérez, H., & Zechinelli-Martini, J.-L. (2017). Humanitarian Logistics and Cultural Diversity within Crowd Simulation. *Computación y Sistemas*, 7-21.

Oji, S., & Adamu, H. (2021). Air Pollution Exposure Mapping by GIS in Kano Metropolitan Area. *Pollution*, 101-112.

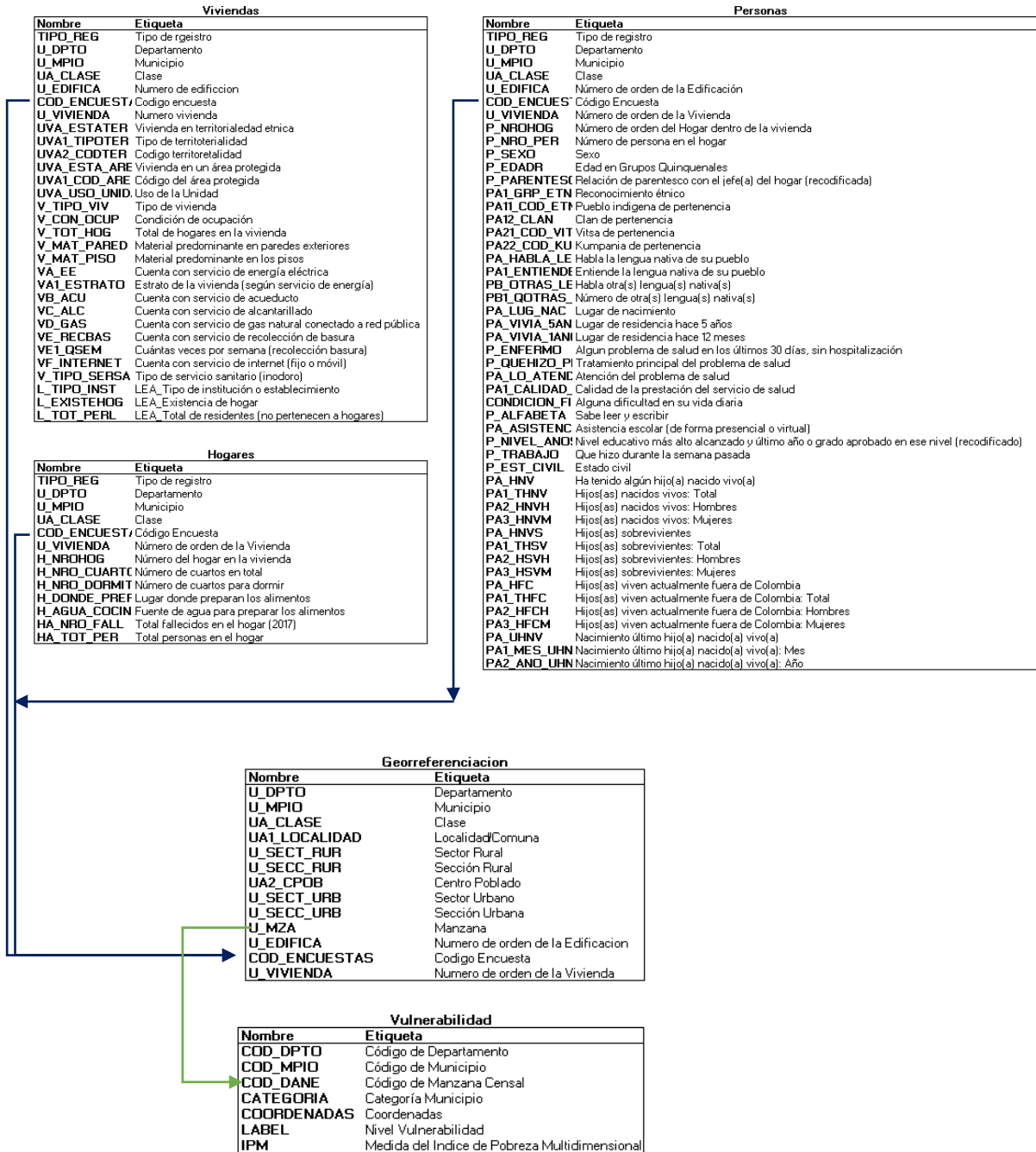
OPS. (2000). *Manual para el Manejo Logístico de Suministros Humanitarios*. Washington, D.C.

Ortiz, D. (2020). Geographic Information Systems (GIS) in Humanitarian Assistance: A Meta-Analysis. *A Journal of Humanistic and Social Inquiry*.

- Qureshi, A. G., & Taniguchi, E. (2020). A multi-period humanitarian logistics model considering limited resources and network availability. *Transportation Research Procedia*, 212-219.
- Rahman, M., Ningsheng, C., Mahmud, G. I., Islam, M. M., & Pourghasemi, H. R. (2021). Flooding and its relationship with land cover change, population growth, and road density. *Geoscience Frontiers*.
- Ramirez, C. O. (2016). *Mecanismos de coordinación para la optimización del desempeño de la cadena logística humanitaria mediante modelamiento estocástico. Caso colombiano*. Bogota.
- Rodríguez-Espíndola, O., Albores, P., & Brewster, C. (2018). Disaster preparedness in humanitarian logistics: A collaborative approach for resource management in floods. *European Journal of Operational Research*, 978-993.
- Scarpin, M. R., & Silva, R. d. (2014). Humanitarian Logistics: Empirical Evidences from a Natural Disaster. *Procedia Engineering*, 102-111.
- Tacheva, Z., & Simpson, N. (2019). Social network analysis in humanitarian logistics research. *Journal of Humanitarian Logistics and Supply Chain Management*, 492-514.

ANEXOS

• Anexo 1: Modelo de datos relacional



- **Anexo 2: Modelo de datos Geoespacial**

Seguridad

Nombre	Etiqueta
SEXO	sexo de la víctima, o cuando lo que se mide es la comisión de un delito del presunto indiciado
EDAD	edad de la víctima, o cuando lo que se mide es la comisión de un delito, será del presunto indiciado
ESTADO	estado civil de la víctima, o cuando lo que se mide es la comisión de un delito, será del presunto indiciado
GRUPO_U	de la víctima, o cuando lo que se mide es la comisión de un delito, será del presunto indiciado. El grupo actor indica el grupo ej. Guerrilla, Paramilitares etc
ACTIVIDAD	actividad ilegal a la que se dedica la víctima o el presunto indiciado. Mirar en contexto con conducta
PARENTI	parentesco entre la víctima y el presunto indiciado
OCCUPACI	ocupación de la víctima o del presunto indiciado, Mirar en contexto con conducta
DISCAPA	condición de movilidad reducida de la víctima o indiciado
GRUPO_I	indica grupos especiales, que están asociados de por vida a la persona y que hacen que sean más vulnerables a un hecho en particular
MEDIO_T	medio transporte donde se movilizaba la víctima o el presunto indiciado según el caso
NIVEL_A	nivel académico de la víctima o el presunto indiciado según el caso
TESTIGC	indica si hubo o no un testigo de los hechos
NOMBRE	nombre del barrio donde ocurrieron los hechos
CODIGO_B	código del barrio donde ocurrieron los hechos
CODIGO_C	código de la comuna donde ocurrieron los hechos
LUGAR	lugar donde ocurrieron los hechos. El lugar es una tipificación del urbanismo más cercano al hecho
SEDE_RE	entidad que recepcionó el caso
CONDUCC	es el inventario de indicadores o hechos asociados a la seguridad y convivencia
MODALID	es la forma como se materializa el hecho
CARACTI	variable que permite analizar y tipificar la conducta
CONDUCC	una forma de caracterización adicional
ARMA_M	es el arma, medio o mecanismo con el que se comete el hecho
ARTICUL	artículos del código penal o del código de contravenciones de Policía
CATEGO	caracterización de algunos tipos penales
BIEN	nombre del bien
CATEGO	agrupación de bienes, nivel intermedio de agregación
GRUPO_I	agrupación de bienes, nivel alto de agregación
MODELO	modelo del bien
COLOR	color del bien
PERMISC	tipo de permiso de porte que tiene el bien
UNIDAD_U	unidad del bien
FECHA_F	fecha en que ocurrió el hecho
LATITUD	latitud geográfica sistema de coordenadas wgs84
CASO	1: indicará el caso, 0: podría indicar bienes o persona, dependiendo de la conducta que se busca analizar, es decir si el hecho es contra la vida y la integridad personal
VALOR	valor de uso del bien
CANTIDA	cantidad que debe ser entendida en el contexto de la unidad de medida

Riesgos Naturales

Nombre	Etiqueta
OBJECTID	ID
Codigo	Codigo riesgo
NOMBRE	Nombre zona
Riesgo	Tipo de riesgo
Descripcion	Descripción riesgo
Tipo_Amenaza	Tipo de amenaza
SHAPE_AREA	Datos poligono
SHAPE_LEN	Datos poligono
CATEGORIA_PROTEC	Categoría proteccion
SUBCATEGORIA_PRO	Descripción

Espacio publico

Nombre	Etiqueta
Nombre	Nombre espacio publico
Shape	Datos poligono
Funcion	Uso
CBML	
Categoria	Categoría de espaci publico
Orden	Orden
Dominio	Dominio
Nivel	Nivel
SubCategoria	Sub categoría
COD_COMUNA	Código comuna
NOM_BARRIO	Nombre de barrio

Datos consolidados a nivel de relaciones con coordenadas

