

Apoyar en el estudio, desarrollo e implementación de software para análisis de datos y funciones cartográficas usando las herramientas dispuestas

Daniel Alejandro Celis Calderón

161218217

Universidad de Cundinamarca

Facultad de ingeniería

programa ingeniería de sistemas

Sede Fusagasugá

2022

Apoyar en el estudio, desarrollo e implementación de software para análisis de datos y funciones cartográficas usando las herramientas dispuestas

Trabajo de Grado presentado como requisito parcial para la obtención del título de Ingeniero de Sistemas

Daniel Alejandro Celis Calderón

161218217

DIRECTOR:

Ana Esperanza Merchán Hernández

Universidad de Cundinamarca

Facultad de ingeniería

programa ingeniería de sistemas

Sede Fusagasugá

2022

Resumen

En el presente documento se dará a conocer información acerca del periodo de pasantía en el Observatorio socioeconómico de la Alcaldía de Fusagasugá. Se expondrá las diferentes actividades que se realizaron a lo largo del ciclo de duración de la pasantía, además de esto, se manifestara la problemática con el estancamiento de información en los diferentes documentos en los que se almacena la información a las que se tiene acceso en la entidad, se explicara de qué forma se logró exhibir la información de interés público por medio de diferentes estadísticas y mapas geoestadísticos, que herramientas fueron usadas a lo largo del proceso para la gestión de la información. Se expresará el contenido de cada uno de los diferentes documentos que almacenan la información de las dependencias de la alcaldía, que funcionaron como pilar para la creación de los diferentes medios de manifestación de la información, explicando el por qué se usaron esos datos específicos en el diseño de las estadísticas y dando una razón del porque no se integraron ciertos datos.

Abstract

This document will provide information about the internship period in the socioeconomic observatory of the mayor's office of Fusagasugá. The different activities that were carried out throughout the internship cycle will be presented, in addition to this, the problem with the stagnation of information in the different databases to which the entity has access will be shown, it will be explained how the information of public interest was displayed through different statistics, The content of each of the different databases that were used as the basis for the creation of the different means of displaying the information, explaining why specific data were used in the design of the statistics and giving a reason why certain data were not integrated.

Glosario

Dashboard: Herramienta de gestión de información, expone, gestiona y analiza de manera visual indicadores de información, para la realización de un seguimiento y estudio por parte de una organización. (Portilla - 2021).

PII (información de identificación personal): Datos que pueden ser usados para identificar de manera directa a un individuo, consta de información de contacto, identificación o localización de una persona. (SAS-2022).

Tableau: Software para diseño y exposición de datos interactivos, por medio del uso de bases de datos, que optimiza el proceso de gestión de datos, para la identificación de patrones y tendencias en los datos. (Baque - 2015)

Georreferenciación: Consta de una técnica de posicionamiento especial, por medio de un sistema de coordenadas (Latitud y longitud). (Camacho & Dávila - 2012).

Observatorio socioeconómico: realiza la búsqueda, la recopilación, el procesamiento, la investigación y la publicación de indicadores sociales, de género, económicos, demográficos, seguridad y convivencia. (Observatorio Socioeconómico Fusagasugá, 2020).

Diagnostico sectorial: Proceso que permite analizar el estado actual de un sector a través de la recopilación de información, interpretación y obtención de conclusiones. (contraloría de Bogotá D.C.- 2022).

Calidad de datos: hace referencia a la conformidad que los datos presentan acorde con el propósito para el cual son designados, por lo que las organizaciones bajo este concepto analizan si estos son confiables, consistentes, si se encuentran actualizados y si están libre de duplicidades (Magal & Rincón - 2019).

Tabla de contenido

Introducción.....	13
Planteamiento del problema.....	14
Formulación del problema.....	15
Árbol de problemas.....	16
Objetivos.....	17
Objetivo general.....	17
Objetivos específicos.....	17
Justificación.....	18
Marco referencial.....	19
Antecedentes de estudio.....	19
Marco teórico.....	22
Marco Legal.....	28
Marco metodológico.....	31
Metodología CRISP-DM (Cross Industry Standard Process for Data Mining).....	31
Plan de proyecto.....	33
Resultados.....	34
Entendimiento del negocio.....	34
Comprensión de problema.....	34
Determinación de objetivos.....	36
Evaluación de la situación actual.....	37

Diagrama de procesos	39
Definición de requerimientos del proyecto	40
Historias de usuario	40
Cronograma de actividades	46
Entendimiento de los datos	47
Proceso de análisis, filtración de datos para el diseño de los Dashboard	47
Información proporcionada por la organización (formatos y medio de envío)	47
Proceso de análisis e interpretación de los datos almacenados	48
Análisis de la información almacenada (Manejo de Google Colab)	49
Preparación de los datos	56
Selección, conversión y corrección de datos para el Dashboard.....	56
Selección de datos.....	58
Proceso para obtener la latitud y longitud con base a las direcciones.....	61
Diseño de las estadísticas en Tableau.....	66
Apartados desarrollados con las herramientas disponibles dentro de Tableau	72
Etiquetas.....	73
Desarrollo de mapas de georreferenciación de la información.....	74
Implementación de filtros de información	76
Creación de los modelos	78
Creación del Dashboard	78
Evaluación y pruebas	85

Correcciones, evaluación y pruebas	85
Análisis de concordancia de datos	85
Verificación de filtros y etiquetas	86
Cambios requeridos por el director de proyecto	86
Despliegue o puesta en producción.....	90
Productos desarrollados	94
Recomendaciones.....	97
Conclusiones.....	98
Referencias bibliográficas.....	100

Tabla de ilustraciones

Ilustración 1.....	16
Ilustración 2.....	25
Ilustración 3.....	26
Ilustración 4.....	27
Ilustración 5.....	27
Ilustración 6.....	31
Ilustración 7.....	38
Ilustración 8.....	39
Ilustración 9.....	48
Ilustración 10.....	50
Ilustración 11.....	51
Ilustración 12.....	52
Ilustración 13.....	53
Ilustración 14.....	54
Ilustración 15.....	55
Ilustración 16.....	56
Ilustración 17.....	58
Ilustración 18.....	62
Ilustración 19.....	62
Ilustración 20.....	64

Ilustración 21.....	65
Ilustración 22.....	66
Ilustración 23.....	67
Ilustración 24.....	68
Ilustración 25.....	69
Ilustración 26.....	69
Ilustración 27.....	70
Ilustración 28.....	71
Ilustración 29.....	72
Ilustración 30.....	73
Ilustración 31.....	73
Ilustración 32.....	74
Ilustración 33.....	75
Ilustración 34.....	76
Ilustración 35.....	77
Ilustración 36.....	77
Ilustración 37.....	78
Ilustración 38.....	80
Ilustración 39.....	82
Ilustración 40.....	83
Ilustración 41.....	84

Ilustración 42.....	87
Ilustración 43.....	89
Ilustración 44.....	90
Ilustración 45.....	91
Ilustración 46.....	92
Ilustración 47.....	93
Ilustración 48.....	93

Índice de tablas

Tabla 1	33
Tabla 2	40
Tabla 3	41
Tabla 4	41
Tabla 5	42
Tabla 6	42
Tabla 7	43
Tabla 8	43
Tabla 9	44
Tabla 10	44
Tabla 11	45
Tabla 12	46
Tabla 13	59

Introducción

Se comprende que la información, está constituida por un grupo de datos supervisados y ordenados, de tal manera que, como finalidad se dé un mensaje que le otorgue sentido y significado a algo real; para llevar a cabo esta acción se requiere que los datos recolectados sean sometidos a un análisis de calidad exhaustivo, que le permita a un sector determinado realizar el debido manejo de la información, ya que para llevar a cabo la ejecución de un proceso estadístico, la gestión tiene la opción de apoyarse en el uso de las tecnologías de la información (TI) la cual brinda una combinación de medios y métodos de recopilación, procesamiento y transmisión de datos para obtener nueva información de calidad.

El presente documento expondrá los procedimientos realizados a un conjunto de datos almacenados en documentos de formato .xlsx (Excel), facilitadas por las dependencias del observatorio socioeconómico de la alcaldía de Fusagasugá.

Dentro de los procesos realizados para este proyecto, se encuentra la transformación de los datos de archivo plano a diseño estadístico (graficas de barras verticales, barras apiladas, mixtas, circulares, burbujas agrupadas y de anillo) y modelos de exposición de información por medios geográficos (Mapas de calor, mapas de comuna y corregimiento y mapas de división sectorial). El diseño de estas estadísticas y mapas de información, facilitaron la exposición información sectorial, creando un medio de interpretación de información entendible para un usuario del común, pero más importante, para la interpretación y análisis de datos por parte de algunas de las ramas de las áreas presentes dentro de la alcaldía de Fusagasugá (secretaria de salud, movilidad, desarrollo económico, entre otras).

Planteamiento del problema

En la alcaldía de Fusagasugá existe una carencia de medios para organizar y presentar información social y económica oportuna, lo que dificulta la toma de decisiones relevantes de la administración central y las ramas que contribuyen en esta, esto simultáneamente provoca, dificultades en la generación de información para la comunidad bajo parámetros de pertinencia.

El proyecto de implementación de un observatorio socioeconómico dentro de la alcaldía de Fusagasugá nace de la intención de solucionar una situación con efectos negativos resultante de la falta de medios para la gestión y organización de información, ya que es una instancia que se encarga de realizar la recopilación, el procesamiento, la investigación y la publicación de indicadores sociales, de género, económicos, demográficos, seguridad y convivencia.

El problema central corresponde a la carencia de medios de manifestación de información proporcionada por las dependencias para el análisis, veraz y oportuno por parte de las oficinas de la alcaldía (para la formulación de diagnósticos sectoriales) y usuarios comunes, se hace más fácil identificar las Causas que provocan dicha situación de forma directa e indirectamente. A nivel de Causas Directas se plantean dos, las cuales son hechos concretos que generan el problema central:

- Estancamiento de la información almacenada dentro de la entidad.
- Inexistencia de herramientas o medios para la gestión y presentación de datos almacenados.

Por otro lado, a continuación, se relacionan las Causas Indirectas las cuales son acciones o hechos que dan origen a las causas directas:

- Falta de instrumentos de procesamiento de información.
- Descentralización de la información estadística.
- Falta de asignación de responsables del manejo de la información estadística dentro del observatorio.

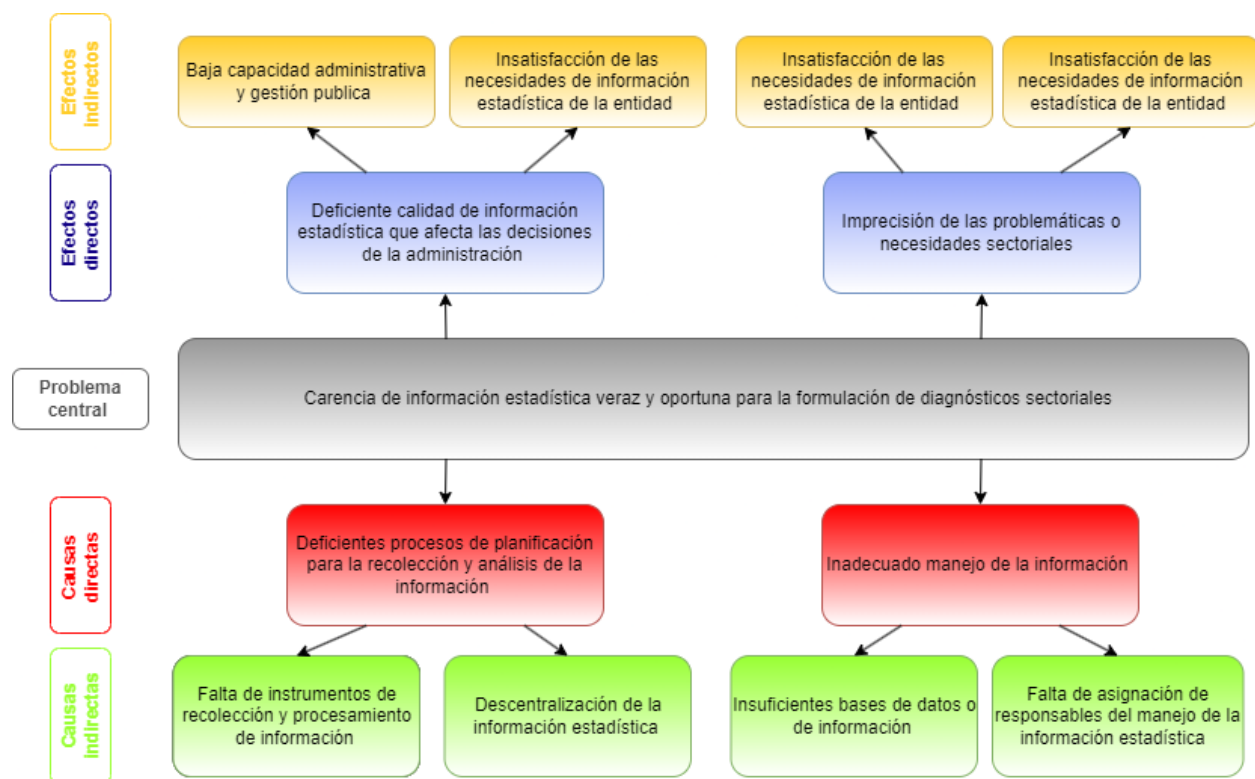
Por lo antes expuesto, es necesario adaptar la mayor cantidad de información posible, una vez se disponga de esta, se debe realizar un estudio de los datos que se tiene a disposición, para identificar la información de mayor relevancia y de mayor necesidad. Con el análisis de la información a disposición se puede dar pie a la creación de las diferentes graficas estadísticas (Medio escogido para exponer la información), que se requieren para la página web de la Alcaldía de Fusagasugá, las cuales desglosaran información de componente geográfico, realizar estadísticas sectoriales de la ciudad y así publicar la información relevante en un Dashboard, facilitando interpretación de información de usuarios comunes que transiten en la página web del observatorio (Los productos finales serán publicados en la página web del observatorio), o para el análisis, estudios o tomas de decisiones por parte de los diferentes órganos que hacen parte del ente.

Formulación del problema

¿Es probable, que, la conversión de datos planos a información estadística mejore el análisis y comprensión de información por parte de las áreas internas de la alcaldía y visitantes de la página web del observatorio?

Árbol de problemas

Ilustración 1.
Árbol de problemas.



Fuente: propia.

Objetivos

Objetivo general

Garantizar el acceso a información estadística veraz y oportuna para la formulación de diagnósticos sectoriales.

Objetivos específicos

- Adecuar la presentación de los datos estadísticos que se manejan, con ayuda de una interfaz gráfica.
- Mostrar los resultados del análisis de los datos, en los mapas de georreferenciación.
- Categorizar los resultados estadísticos por sector, dependiendo de los datos, en los mapas de georreferenciación.

Justificación

La estadística es una ciencia derivada de las matemáticas, expone información de forma ordenada y entendible, la cual es aplicada para la recolección y análisis de datos en estudio, con el objetivo principal de estimar resultados que ayuden a las organizaciones con la toma de decisiones.

En vista de que dentro de la alcaldía de Fusagasugá, hasta hace poco, no se contaba con un área que se encargara de la gestión información sectorial, por esta razón, Se realizo un proyecto para implementación de una nueva rama dentro de la estructura organizacional de la alcaldía (Observatorio socio económico de la alcaldía de Fusagasugá), esta área actualmente se encarga de realizar la búsqueda, la recopilación, el procesamiento, la investigación y la publicación de indicadores sociales, categorizados por género, económicos, demográficos, de seguridad o de convivencia.

La proyección que se planteó en el área estadística dentro del observatorio para gestión de información concluyo en el uso de gráficos estadísticos sectoriales y el uso de mapas geoestadísticos para mostrar la información una vez procesada (seleccionada, filtrada y transformada). Estas estadísticas y mapas antes mencionados fueron almacenados en un Dashboard (cuadro de mando), para la organización y representación graficas de estos componentes, una vez completado el proceso de construcción del producto final, se realizaba la publicación por medio de la página web oficial asignada al Observatorio, para que así tanto usuarios como áreas de la alcaldía pudieran hacer un estudio, análisis y visualización de los datos (en el caso de las áreas de la alcaldía, ayudo en la toma de decisiones, análisis e interpretación de información).

Marco referencial

Antecedentes de estudio

Para la construcción de este documento, se estudiaron una serie de documento de proyectos realizados con temáticas similares a las que se propone en la realización de este trabajo, estos documentos antes mencionados proveen de una serie de temáticas, metodologías de trabajo y conceptos, que aportaran en el desarrollo de este proyecto.

El autor del proyecto “Análisis de la base de datos del hospital universitario de guayaquil para la creación de reportes gráficos con la herramienta Tableau”, propone un proyecto de implementación dentro de una entidad hospitalaria, usando como base la herramienta de Tableau Public mejorando el manejo de consultas e informando a las personas diferentes síntomas que pueden padecer, todo enfocándose a una enfermedad que afecta a las rodillas de los individuos que la padecen. Enfocándose en la prevención de la osteoartritis de rodilla, tratarla y dar una mejor consulta y advertir a las personas que tiene sobre peso se haga chequeo para prevenir a futuro los dolores articulares en la rodilla. (Baque ,2014).

La autora del proyecto “Análisis de los pacientes ingresados en ucín vs el número de nacimientos vivos del hospital universitario de guayaquil, para la generación de reportes utilizando la herramienta Tableau Public y sql server 2008.” Desarrollo una tesis con la finalidad de proveer el resultado detallado del análisis de datos en el hospital Universitario de Guayaquil, ya que, por falta del estudio, se estaban presentando problemas de retraso de la información que desean visualizar los médicos especialistas encargados de dar seguimiento a pacientes. (Villavicencio - 2015).

El autor del proyecto “Desarrollar un conjunto de tableros visuales de Business Intelligence (BI) mediante la herramienta de Tableau que permita fortalecer la interpretación de la información de deserción estudiantil en la Universidad Técnica del Norte” desarrolla un proyecto de tesis con el propósito de fortalecer el avance tecnológico y ser de referencia para futuras investigaciones que propongan el uso de BI en los negocios de la zona, investigando aportaciones anteriores validas y de calidad que serán de utilidad para la aplicación del proyecto que desarrollo (Silva - 2022).

El autor del proyecto “Calidad de datos”, comenta la necesidad e importancia del manejo de datos dentro de un proyecto, aporta una serie de interrogantes que resuelve a lo largo del desarrollo del documento propuesto, estas inquietudes antes mencionadas constan de, qué aporta un dato, cómo implementar la calidad de datos, qué debemos se debe tener en cuenta como mejores prácticas y qué tecnologías soportan la calidad de los datos. (Cabanillas – N.d.).

El autor del proyecto “Aplicación de metodología CRISP-DM para segmentación geográfica de una base de datos pública”, propone la implementación la metodología CRISP-DM para solventar la falta de medios para la extracción de información útil para la solución de problemas reales, el autor desarrollo un modelo de segmentación geográfica, con ayuda del lenguaje de programación R para el análisis estadístico y de distribución. (Espinosa - 2020).

El autor del proyecto “Aplicación de la metodología CRISP-DM a la recolección y análisis de datos georreferenciados desde Twitter”, desarrollo un proyecto con el objetivo de recolectar, clasificar, depurar y almacenar datos provenientes de una aplicación de red social, con el fin de identificar información en un espacio geográfico definido (Bogotá), obteniendo información implícita u oculta, previamente desconocida y potencialmente útil de grandes volúmenes de información. (García - 2018).

El autor del proyecto “Conector web para la integración de datos abiertos colombianos con la herramienta Tableau”, desarrollo un aplicativo web en donde se interactúa con datos abiertos del portal gubernamental de “Datos abiertos para Colombia”, este aplicativo funciona de la mano con el programa de manejo de datos Tableau, permitiendo personalizar la información que será visualizada por las personas o entidades interesadas en analizar el comportamiento de los diferentes indicadores que afectan una ciudad. (Gómez - 2018).

La autora del proyecto “Estandarización de los reportes del análisis de resultados de las pautas publicitarias digitales de la agencia MediaCom Colombia para sus clientes a través del software Tableau “agrupa información proveniente de la agencia MediaCom, se determinó que era necesario estandarizar un informe general para la gestión de información, para hacer entrega de esta a los diferentes clientes que intervienen en la empresa. (Escobar - 2016).

Marco teórico

Para la realización del proyecto se debe tener en cuenta una serie de términos, para lograr entender los diferentes procesos realizados a lo largo de la pasantía, se dará explicación con fundamentación teórica de las diferentes terminologías que intervienen con el proceso de diseño de estadísticas, análisis de datos, filtración de datos, creación de coordenadas a partir de información geo referente de un sitio, diseño de mapas a partir de información registrada, y diseño de Dashboard para la presentación e información.

Georreferenciación

Es un proceso que permite determinar la posición de un elemento o de un lugar específico, esto por medio del uso de coordenadas especiales, este proceso se determina por medio de la relación entre elementos espaciales de un sistema, de manera que, teniendo información específica de un sistema se puede determinar el origen del sistema de destino, y así poder determinar la posición cartográfica por medio de relación de información vectorial. (Camacho & Dávila - 2012)

Estadística

La estadística es una ciencia que se ocupa de la recolección, agrupación, presentación, análisis e interpretación de datos, su objetivo es describir el conjunto de datos obtenidos y tomar decisiones o realizar generalizaciones acerca de las características de todas las posibles observaciones bajo consideración. (Velázquez, N.d.).

Metodología CRISP-DM

La metodología CRISP-DM (Cross-Industry Standard Process for Data Mining) consta de una serie de actividades para administración de proyectos orientados al manejo de datos. Esta metodología se puede aplicar a una gran variedad de problemas tales como encontrar perfiles de clientes fraudulentos, estimar la probabilidad de que los clientes de una empresa se vayan a la competencia o también determinar patrones de compra para luego, recomendar productos de interés.

CRISP-DM está compuesta por seis fases, las cuales dependen entre sí tanto en forma secuencial como cíclica, pudiendo encontrarse interacciones que permitan mejorar la aproximación obtenida en otras fases anteriores.

Las fases que componen la metodología son las siguientes:

- Comprensión del negocio.
- Comprensión de los datos.
- Preparación de los datos.
- Modelado.
- Evaluación.
- Despliegue.

(Galán - 2015).

Dashborad (Cuadro de mando).

Un Dashboard, también llamado cuadro de mando o panel de control es una herramienta para la gestión de la información, se trata de, una representación gráfica de las de los datos fundamentales que intervienen para hacer el seguimiento de la empresa o de una campaña específica. Es una pantalla que ofrece una representación visual de los datos más relevantes para ofrecer una visión general. (Portilla - 2021).

Calidad de datos

La calidad de datos se refiere a la conformidad de una serie de datos con respecto al propósito con el que fueron designados, identificando si cumplen o no con una serie de requisitos, como lo son:

- **Compleitud:** Verificación de que el medio de almacenamiento no cuente con datos nulos dentro de sus registros.
- **Conformidad:** Los datos deben contar con un formato estándar y legible para todos sus registros.
- **Consistencia:** Evitar el uso de información contradictoria
- **Precisión/exactitud:** Como el nombre lo menciona se trata de verificar la veracidad de un dato, comparándolo con otra fuente o registro.
- **Duplicación:** Evitar el registro de datos repetidos dentro de los registros disponibles.
- **Integridad:** Evitar la modificación de información, cuando no es necesario, ya que puede perder su probidad.

(Magal & Rincón - 2019).

Herramientas

Excel

Microsoft Excel es una herramienta que permite realizar tareas contables por medio del uso de hojas de trabajo, permitiendo manipular los datos almacenados dentro de las hojas, por medio de filas y columnas.

Los medios de almacenamiento obtenidos por parte de cada una de las dependencias (secretarías, oficinas, organismos de control, entre otras), provenían de documento de extensión .xlsx (documentos Excel), por lo que, el inicio de las actividades para llegar al producto final comenzó por medio de una gestión de datos en este programa.

Ilustración 2.

Logo de Excel.



Tomado de: 1000 marcas. Microsoft Excel logo. <https://1000marcas.net/microsoft-excel-logo/>

Tableau

Tableau es una herramienta de visualización de datos utilizada mayormente en el área empresarial, simplifica los datos en bruto en un formato más fácil de entender, ayudar a los usuarios a ver y comprender toda la información que poseen. (Villavicencio - 2015).

Esta herramienta fue un pilar fundamental para la construcción y exposición gráfica de la información, ya que por medio del uso de sus diferentes herramientas (estadísticas, etiquetas, filtros, mapas, entre otras), se logró componer los diferentes productos que serán presentados más adelante.

Ilustración 3.
Logo tableau.



Tomado de: Logos World. Tableau Logo. <https://logos-world.net/tableau-logo/>

Google Colab

Google Colab consta de una plataforma en línea la cual permite escribir y ejecutar bloques de código en el lenguaje Python.

Dentro de los documentos almacenados en el observatorio era necesario tener un medio que mostrara el análisis de la información de los documentos que contenían la información con la que se estuviera trabajando, por lo que, se hizo uso de la herramienta Google Colab, para construir un documento que analice hasta que punto se cumple las dimensiones de la calidad de los datos (completitud, conformidad, consistencia, precisión / exactitud, duplicación e integridad) y a su vez resuma datos de importancia para el observatorio.

Ilustración 4.
Logo Colab.



Tomado de: Wikimedia Commons. Google Colaboratory SVG Logo.svg.

https://commons.wikimedia.org/wiki/File:Google_Colaboratory_SVG_Logo.svg

Google Sheets

Herramienta de manejo de hojas de cálculo, como operaciones muy similares a las realizadas en Microsoft Excel, con diferencia en la cantidad de herramientas que se tienen a disposición.

Este aplicativo web, que fue usado para la transformación de los datos de direcciones almacenados en los documentos proporcionados por las dependencias de la alcaldía de Fusagasugá, usando bloques de código en lenguaje Java Script dentro del apartado de App Script.

Ilustración 5.
Logo Colab.



Google Sheets

Tomado de: ResearchGate. Logo Google Sheets. https://www.researchgate.net/figure/Google-Sheets-TM-Logo_fig1_345911677

Marco Legal

Artículo 2° de la Constitución Política que establece: Son fines esenciales del Estado: servir a la comunidad, promover la prosperidad general y garantizar la efectividad de los principios, derechos y deberes consagrados en la Constitución; facilitar la participación de todos en las decisiones que los afectan y en la vida económica, política, administrativa y cultural de la Nación; defender la independencia nacional, mantener la integridad territorial y asegurar la convivencia pacífica y la vigencia de un orden justo.

Artículo 20 de la Constitución Política determina que: Se garantiza a toda persona la libertad de expresar y difundir su pensamiento y opiniones, la de informar y recibir información veraz e imparcial, y la de fundar medios de comunicación masiva. Estos son libres y tienen responsabilidad social. Se garantiza el derecho a la rectificación en condiciones de equidad. No habrá censura.

Constitución política en su artículo 23 reitera que: Toda persona tiene derecho a presentar peticiones respetuosas a las autoridades por motivos de interés general o particular y a obtener pronta resolución. El legislador podrá reglamentar su ejercicio ante organizaciones privadas para garantizar los derechos fundamentales. De igual manera, el **artículo 74** ratifica que: Todas las personas tienen derecho a acceder a los documentos públicos, salvo los casos que establezca la Ley.

Ley 134 de 1994 en su artículo 1° “Objeto de la Ley”, establece las normas fundamentales por las que se regirá la participación democrática de las organizaciones civiles.

Ley 136 de 1994 en su artículo 5° “Principios Rectores de la Administración Municipal, determina la organización y el funcionamiento de los municipios se desarrollará con arreglo a los postulados que rigen la función administrativa y regulan la conducta de los servidores públicos, y en especial; con sujeción a los principios de eficacia, eficiencia, publicidad y transparencia, moralidad, responsabilidad e imparcialidad.

Ley 489 de 1998 en su artículo 32 “Democratización de la Administración Pública. <Artículo modificado por el artículo 78 de la Ley 1474 de 2011> refiere que todas las entidades y organismos de la Administración Pública tienen la obligación de desarrollar su gestión acorde con los principios de democracia participativa y democratización de la gestión pública. Para ello podrán realizar todas las acciones necesarias con el objeto de involucrar a los ciudadanos y organizaciones de la sociedad civil en la formulación, ejecución, control y evaluación de la gestión pública.

Ley 1551 de 2012 en su artículo 1° “Objeto de la Ley” manifiesta que la presente Ley tiene por objeto modernizar la normativa relacionada con el régimen municipal, dentro de la autonomía que reconoce a los municipios la Constitución y la Ley, como instrumento de gestión para cumplir sus competencias y funciones.

Ley 1712 del 2014 en su artículo 1° “Objeto” refiere que: El objeto de la presente Ley es regular el derecho de acceso a la información pública, los procedimientos para el ejercicio y garantía del derecho y las excepciones a la publicidad de información.

el **decreto 103 de 2015** en su Artículo 1° “Objeto” establece que el objeto es reglamentar la Ley 1712 de 2014, en lo relativo a la gestión de la información pública.

Ley 1753 de 2015 Art 160, Se adopta el Plan Nacional de Desarrollo 2014-2018 “Todos por un nuevo país” en que se determina la creación del Sistema Estadístico Nacional.

Decreto Nacional 1743 de 2016 Plan Estadístico Nacional

La **Directiva 002 de la Procuraduría General de la Nación** del 3 de febrero de 2020.

Establece que los entes territoriales, conforme lo dispuesto en la **Ley 1361 de 2009**, deben establecer acciones, planes y programas tendientes a promover una cultura de protección, promoción y realce de la institución familiar (**Artículo 7°**); establecer Observatorios de familia regionales adscritos a las oficinas de planeación departamentales o municipales (**Artículo 9°**) y formular y ejecutar política pública de apoyo y fortalecimiento de la familia. (**Artículo 13**).

Marco metodológico

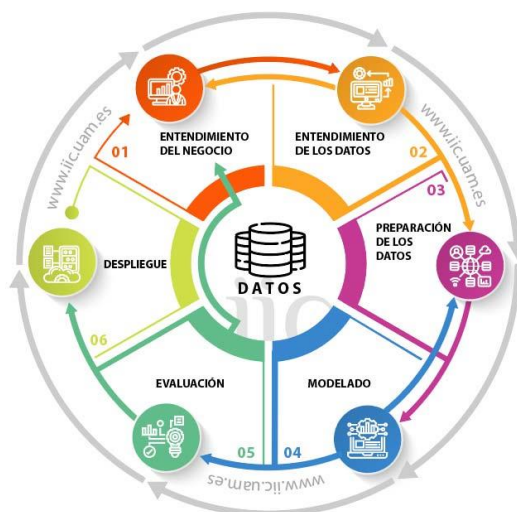
Metodología CRISP-DM (Cross Industry Standard Process for Data Mining)

es una metodología desarrollada específicamente para diseñar proyectos de IT (tecnologías de la información) buscando un mayor control y unos mejores resultados, esta metodología tiene un enfoque hacia el manejo de datos. Esta es la razón por la que fue implementada dentro de proyecto de construcción de estadísticas.

Por medio de esta metodología se gestionó los procesos para el diseño de los Dashboard para la página de la alcaldía, ya que, cada una de sus fases se acoplo al manejo y desarrollo de las actividades necesarias para llevar a cabo la construcción de los productos estadísticos, las estepas antes mencionadas constan de los siguientes seis niveles: (Galán - 2015).

Ilustración 6.

Fases de la metodología CRISP-DM



Tomado de: Instituto de ingeniería del conocimiento. Fases de la metodología CRISP-DM. <https://www.iic.uam.es/innovacion/metodologia-crisp-dm-ciencia-de-datos/>

1. **Entendimiento del negocio:** Esta es la etapa más importante, ya debe tener una correcta comprensión del problema, o negocio a identificando una serie de objetivos a cumplir y analizando la situación actual de la empresa, dentro de esta se puede hacer una subdivisión de actividades las cuales son:

Identificación del problema: Consiste en entender y delimitar la problemática, así como identificar los requisitos, supuestos, restricciones y beneficios del proyecto.

Determinación de objetivos: Determinar las metas a lograr al proponer una solución. El objetivo es obtener un patrón de comportamiento de entidad, de acuerdo con los indicadores de información.

Evaluación de la situación actual: Especifica el estado actual antes de implementar la solución propuesta, a fin de tener éxito del proyecto.

2. **Entendimiento de los datos:** Identificar cómo y dónde reposa la información con que cuenta la organización para llevar a cabo el proyecto. Además de, otros ítems como permisos de acceso y formato de los datos (en este caso Excel).
3. **Preparación de los datos:** Se realiza una preparación de los datos para la continuación de los procesos, ya que, las fuentes de información no están preparadas para el diseño de los Dashboard.
4. **Creación de los modelos:** Dentro de esta etapa se define un modelo para los datos, en el caso de este proyecto, se define una estructura para la construcción de un Dashboard que cumpla con el objetivo planteado en un inicio para los datos.
5. **Evaluación y pruebas:** Los modelos seleccionados deben ser testeados y para verificar su precisión y eficacia.
6. **Despliegue o puesta en producción:** En esta esta etapa final se realiza el despliegue del producto resultado del desarrollo de las etapas anteriores, en el caso del proyecto, se trata de la aprobación para la publicación de los Dashboard dentro de la página oficial del Observatorio socioeconómico de la Alcaldía de Fusagasugá.

Plan de proyecto

El desarrollo tiene la finalidad de construir un conjunto de tableros de análisis integral, cuyo objetivo es la visualización de patrones de datos de las dependencias del Observatorio.

Para la estimación de cada tarea se va a implementar una denominación técnica Shirt Sizes (Sugerida por el autor del proyecto Silva, n.d.), el cual se basa en asignar el esfuerzo para realizar una actividad basándose en las tallas de una camiseta en vez de usar números (XS, S, M L, XL). Para cada tamaño se le puede dar un valor numérico una vez finalizada la estimación de las tareas, como ejemplo en la siguiente Tabla 1, se muestra una posible manera de valoración numérica.

Tabla 1
Medidas para las historias de usuario

Estimación	Nombre de estimación	Valor numérico esfuerzo
S	Extrapequeña	5
S	Pequeña	15
M	Mediana	25
L	Grande	50
XL	Extragrande	100

Resultados

Entendimiento del negocio

Iniciando la aplicación de la metodología se dará pie al apartado de entendimiento del negocio dentro de esta etapa, se expondrá la problemática de la entidad (Observatorio de la alcaldía de Fusagasugá), y dando contexto de la entidad.

Comprensión de problema

Dentro de la entidad pública, Alcaldía de Fusagasugá Cundinamarca, no se contaba con medios para mostrar información estadística de los procesos que se llevan a cabo en cada una de las entidades dependientes de orden municipal, generando un estancamiento de información y a su vez generando pérdidas, por motivos, de usar espacios de almacenamiento de información no dinámicos, lo cual causó un bajo conocimiento eficaz y oportuno de los usuarios, ante el manejo y publicación de la información, a raíz de esta necesidad, el área de finanzas públicas y desarrollo económico, implementó un plan de desarrollo, donde se pudiera alojar información dinámica y oportuna de los procesos estadísticos llevados a cabo por las entidades dependientes de la Alcaldía Municipal de Fusagasugá, dando como nombre a esta nueva instancia (Observatorio socioeconómico de la alcaldía de Fusagasugá), dentro de los organismos del ente. En esta nueva rama generada se empezaron a ejecutar actividades encaminadas al diseño de indicadores de información.

Estos procesos son realizados de la mano de las diferentes oficinas que conforman a la empresa, ya que, el proceso de recolección de información es por medio de, (entrevistas, estudios sectoriales, encuestas, entre otros), es realizado por cada oficina dependiendo de su campo de trabajo, por ejemplo un documento Excel que contiene información de natalidad dentro del municipio, es proporcionada por la secretaria de salud de la alcaldía, esta oficina

interna de la alcaldía, deben realizar el análisis de los datos recolectados, ya que estos deben ser datos verídicos y de calidad, para que al llegar al área encargada del observatorio, entren a proceso de filtración, transformación y cargue de los mismos en las herramientas usadas por el proceso estadístico, para su publicación.

Ya que se tiene acceso a una gran cantidad de información proporcionada por cada dependencia, es necesario tener un medio por el cual se pueda mostrar la información, siendo este, claro y de fácil entendimiento tanto para un usuario común como para una persona que realice actividades laborales dentro de la alcaldía, es por esta razón, que se determinó dentro del observatorio que el uso de gráficos estadísticos y mapas de georreferenciación de información, sería un medio viable para dar cumplimiento al requerimiento antes mencionado.

Uno de los mejores medios de socialización de información es por medio gráfico, esto beneficiara en gran medida a los usuarios que deseen o que necesiten analizar la información. La construcción de las estadísticas facilitará los procesos de toma de decisiones de cada una de las oficinas establecidas dentro de la alcaldía, ya que se disminuirá el tiempo y complejidad a la hora de realizar un estudio de información (Beneficiando a las oficinas en la ejecución de actividades).

La construcción de las estadísticas partió de la socialización de los documentos que almacenan la información proporcionada por las dependencias del Observatorio, esto genero una serie de restricciones considerables como lo fueron:

- **Incompletitud de datos:** En cada uno de los documentos proporcionados por las dependencias, se encontraron series de campos nulos (normalmente marcados con símbolos como guiones “-” o barras invertidas “/”), en datos importantes como lo pueden ser edad, dirección, genero entre otros.

- **Inconformidad de datos:** los datos que se encontraban dentro de los campos no contaban con un formato estándar, por lo que cada dependencia contaba con un orden de almacenamiento propio.
- **Inexactitud de datos:** En ciertos campos específicos de las tablas, se encontraban datos imprecisos, ya sea por contener información incompleta o por contar con información inexistente (en el caso de las direcciones almacenadas).
- **Duplicación de información:** dentro de los registros realizados por los usuarios, se encontraban campos repetidos, en casos se presentaba la duplicación tres veces, generando inconsistencias en los resultados finales expuestos dentro de las estadísticas.

Determinación de objetivos

A continuación, se especificará los objetivos a cumplir para dar solución a la problemática expuesta con anterioridad, se busca obtener un patrón de comportamiento de datos de la entidad pública, de acuerdo con los indicadores de información. Dentro del estudio hacia la entidad pública se logró determinar los siguientes objetivos a cumplir:

- Proporcionar un medio de información (gráficas estadísticas) para que las dependencias de la alcaldía de Fusagasugá puedan realizar la verificación de la información proporcionada de forma más ordenada, oportuna y eficaz.
- Establecer un fin para la información almacenada dentro de la alcaldía (suprimir el estancamiento de información)
- Suprimir la descentralización de la información, generando sistema de información por medio de indicadores de información socioeconómicos en base los datos proporcionada por cada dependencia (sistema realizado por el área de desarrollo de software del observatorio socioeconómico).

Evaluación de la situación actual

Especifica el estado actual antes de implementar la solución de gestión de datos propuesta, a fin de tener un punto de comparación que permita medir el grado de éxito del proyecto.

En un inicio dentro de la alcaldía no se contaba con un área asignada para la administración de información, es por ello que, el área de planificación de desarrollo y finanzas Públicas, de la alcaldía de Fusagasugá, presentó un plan de desarrollo “Implementación del Observatorio Socioeconómico en el Municipio de Fusagasugá Dirección de Planificación del Desarrollo y Finanzas Públicas Alcaldía de Fusagasugá, año 2021”, con el cual se pretende que todas las áreas dependientes de la organización, generen sus indicadores socioeconómicos (en los indicadores se manejan estadísticas de población, economía, movilidad, comercio etc.) basados en la recolección de datos y de información sectorial, conforme a los lineamientos que manejen cada una de ellas.

A continuación, se nombran algunas de las dependencias que aportaron datos para su gestión:

- Secretaria de salud.
- Secretaria de movilidad.
- Secretaria de desarrollo económico.
- Secretaria de solidaridad.
- Ordenamiento territorial.

Estructura organizacional de la entidad pública (Observatorio socioeconómico y alcaldía de Fusagasugá).

Ilustración 7.
Marco organizacional

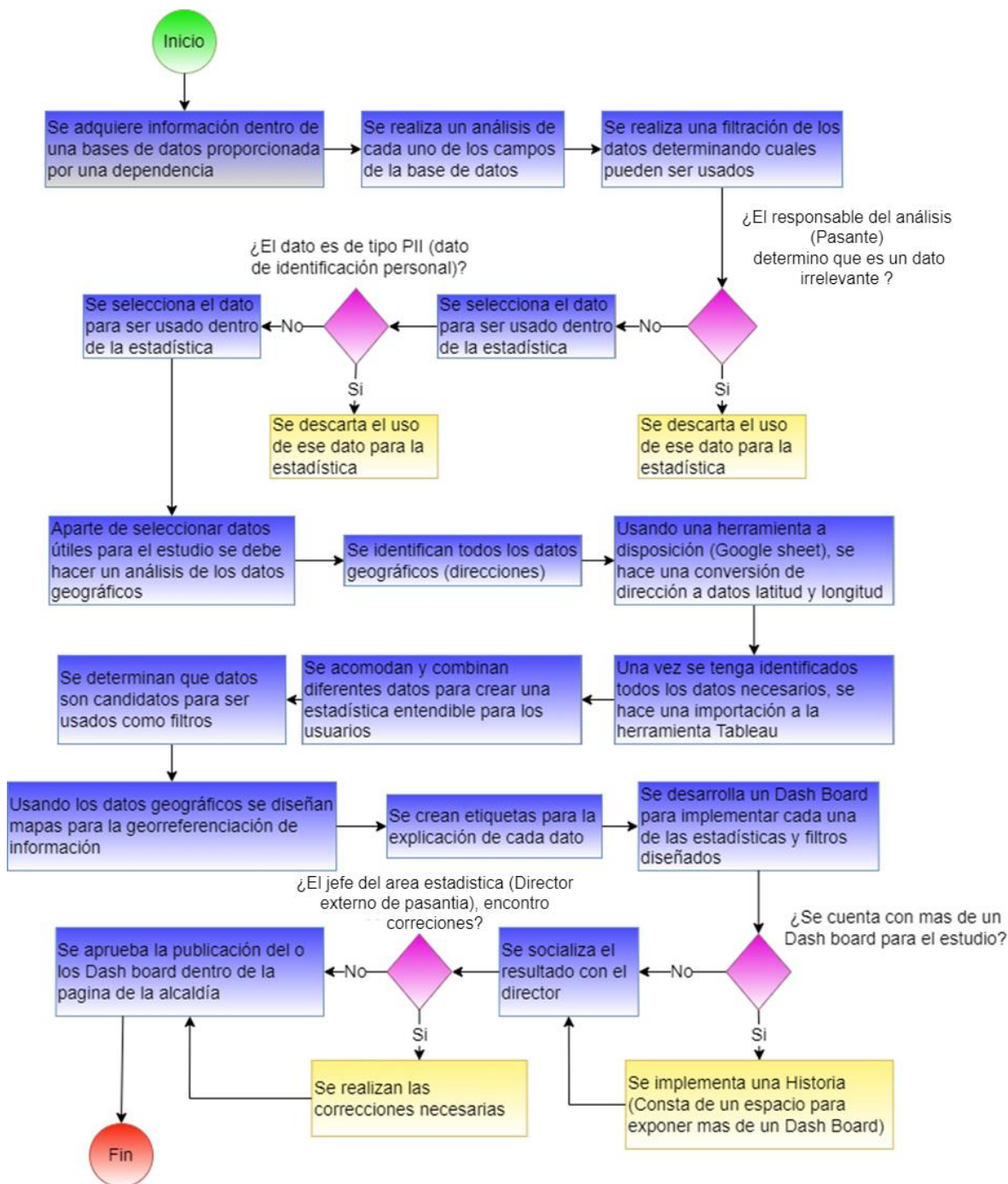


Fuente: Propia

Diagrama de procesos

Ilustración 8.

Diagrama de procesos



Fuente: Propia.

Definición de requerimientos del proyecto

Para la definición de requerimientos se realizaron actividades para determinar las necesidades de los usuarios, y así establecer las historias de usuario. Las actividades constaron de reuniones con el equipo de desarrollo del Observatorio socioeconómico de la Alcaldía, realizando entrevistas para recolectar las variables requeridas para el posterior análisis.

Historias de usuario

De las entrevistas se establecieron las historias de usuario requeridas por el cliente, las cuales se redactaron también en base a las variables de interpretación de datos.

Tabla 2

Historia de usuario natalidad

Historia de usuario	
Numero: H1	Usuario: Desarrollador
Nombre de la historia: Grafico estadístico natalidad	
Prioridad: Alta	Esfuerzo: XL
Descripción: Realizar un tablero estadístico que exponga el porcentaje de los datos de mayor importancia dentro del documento Excel que almacena datos de natalidad, para conocer los porcentajes en la natalidad en general y la natalidad en mujeres menores de edad (14 a 17), además de las cantidades de nacimientos de niños y vivos.	
Prueba de aceptación: Dashboard iterativo relacionado con las natalidades en la ciudad de Fusagasugá.	

Tabla 3

Historia de usuario seguridad y convivencia 2019

Historia de usuario	
Numero: H2	Usuario: Desarrollador
Nombre de la historia: Grafico estadístico seguridad y convivencia 2019	
Prioridad: Alta	Esfuerzo: XL
Descripción: Realizar un tablero estadístico que exponga el porcentaje de los datos de mayor importancia dentro del documento Excel que almacena información de seguridad y convivencia en Fusagasugá del año 2019.	
Prueba de aceptación: Dashboard iterativo relacionado con el reporte de seguridad y convivencia en la ciudad de Fusagasugá para el año 2019.	

Tabla 4

Historia de usuario seguridad y convivencia 2020

Historia de usuario	
Numero: H3	Usuario: Desarrollador
Nombre de la historia: Grafico estadístico seguridad y convivencia 2020	
Prioridad: Alta	Esfuerzo: XL
Descripción: Realizar un tablero estadístico que exponga el porcentaje de los datos de mayor importancia dentro del documento Excel que almacena información de seguridad y convivencia en Fusagasugá del año 2020.	
Prueba de aceptación: Dashboard iterativo relacionado con el reporte de seguridad y convivencia en la ciudad de Fusagasugá para el año 2020.	

Tabla 5
Historia de usuario seguridad y convivencia 2021

Historia de usuario	
Numero: H4	Usuario: Desarrollador
Nombre de la historia: Grafico estadístico seguridad y convivencia 2021	
Prioridad: Alta	Esfuerzo: XL
Descripción: Realizar un tablero estadístico que exponga el porcentaje de los datos de mayor importancia dentro del documento Excel que almacena información de seguridad y convivencia en Fusagasugá del año 2021.	
Prueba de aceptación: Dashboard iterativo relacionado con el reporte de seguridad y convivencia en la ciudad de Fusagasugá para el año 2021.	

Tabla 6
Historia de usuario defunciones 2020

Historia de usuario	
Numero: H5	Usuario: Desarrollador
Nombre de la historia: Grafico estadístico defunciones del año 2021	
Prioridad: Alta	Esfuerzo: XL
Descripción: Realizar un tablero estadístico que exponga las defunciones ocurridas en el año 2021 en la ciudad de Fusagasugá, usando como pilar el documento Excel que almacena información proporcionada por el Observatorio socioeconómico de la Alcaldía de Fusagasugá.	
Prueba de aceptación: Dashboard iterativo relacionado con el reporte de defunciones en la ciudad de Fusagasugá para el año 2021.	

Tabla 7

Historia de usuario seguridad y convivencia 2020

Historia de usuario	
Numero: H6	Usuario: Desarrollador
Nombre de la historia: Grafico estadístico defunciones fetales	
Prioridad: Alta	Esfuerzo: XL
Descripción: Realizar un tablero estadístico que exponga las defunciones, profundizando en las muertes fetales ocurridas en el año 2021 en la ciudad de Fusagasugá, usando como pilar el documento Excel que almacena información proporcionada por el Observatorio socioeconómico de la Alcaldía de Fusagasugá.	
Prueba de aceptación: Dashboard iterativo relacionado con el reporte de defunciones en la ciudad de Fusagasugá para el año 2021.	

Tabla 8

Historia de filtros

Historia de usuario	
Numero: H7	Usuario: Desarrollador
Nombre de la historia: Filtros	
Prioridad: Baja	Esfuerzo: S
Descripción: Dentro de los tableros realizados para la exposición de la información albergada dentro de los documentos que almacenan los datos de la dependencia, integrar una función de filtrado, para modificar las cantidades expuestas dependiendo de la opción seleccionada en el filtro.	
Prueba de aceptación: Herramienta de filtrado de información dentro de lo tableros diseñados.	

Tabla 9
Historia de usuario “historias”

Historia de usuario	
Numero: H8	Usuario: Desarrollador
Nombre de la historia: Historias	
Prioridad: Media	Esfuerzo: M
Descripción: Dentro de los productos realizado integrar el uso de historias, para agrupar más de un Dashboard, y así poder exponer de modo ordenado, dos o más tableros de una misma temática.	
Prueba de aceptación: Sección de historias para seleccionar un tablero diferente dependiendo de la temática manejada.	

Tabla 10
Historia de usuario “historias”

Historia de usuario	
Numero: H9	Usuario: Desarrollador
Nombre de la historia: Gráficos estadísticos violencia intrafamiliar	
Prioridad: Alta	Esfuerzo: XL
Descripción: Realizar un tablero estadístico que exponga los casos de violencia intrafamiliar en la ciudad de Fusagasugá, usando como pilar el documento Excel proporcionado por el Observatorio socioeconómico de la Alcaldía de Fusagasugá.	
Prueba de aceptación: Sección de historias para seleccionar un tablero diferente dependiendo de la temática manejada.	

Tabla 11

Historia de usuario “Tabla de esfuerzo para la realización de requerimientos”

Código	Nombre del requerimiento	Esfuerzo
RF 1	Gráfico estadístico de natalidad	XL
RF 2	Gráfico estadístico seguridad y convivencia 2019	XL
RF 3	Gráfico estadístico seguridad y convivencia 2020	XL
RF 4	Gráfico estadístico seguridad y convivencia 2021	XL
RF 5	Gráfico estadístico defunciones del año 2021	XL
RF 6	Gráfico defunciones fetales	XL
RF 7	Filtros	S
RF 8	Historias	M

Cronograma de actividades

Tabla 12
Cronograma de actividades

Actividades realizadas	MES 1	MES 2	MES 3	MES 4	MES 5	MES 6	MES 7	MES 8
Entendimiento del negocio (Observatorio)								
Entendimiento de los datos (Indicador de nacimientos)								
Preparación de los datos (Indicador de nacimientos)								
Diseño del modelo (Dashboard Indicador de nacimientos)								
Evaluación indicador de nacimientos (Correcciones por parte del director de pasantía)								
Puesta en producción del Dashboard de nacimientos (publicación en la página web del observatorio)								
Entendimiento de los datos (indicador de hechos delictivos)								
Preparación de los datos (indicador de hechos delictivos)								
Diseño del modelo (Dashboard Indicador de hechos delictivos)								
Evaluación indicador de hechos delictivos (Correcciones por parte del director de pasantía)								
Puesta en producción del Dashboard de hechos delictivos (publicación en la página web del observatorio)								
Entendimiento de los datos (indicador de defunciones)								
Preparación de los datos (indicador de defunciones)								
Diseño del modelo (Dashboard indicador de defunciones)								
Evaluación indicador de defunciones (Correcciones por parte del director de pasantía)								
Puesta en producción del Dashboard de defunciones (publicación en la página web del observatorio)								
Entendimiento de los datos (indicador de establecimientos comerciales)								
Preparación de los datos (indicador de establecimientos comerciales)								
Diseño del modelo (Dashboard indicador de establecimientos comerciales)								
Evaluación indicador de establecimientos comerciales (Correcciones por parte del director de pasantía)								
Puesta en producción del Dashboard de establecimientos comerciales (publicación en la página web del observatorio)								
Entendimiento de los datos (indicador de violencia intrafamiliar)								
Preparación de los datos (indicador de violencia intrafamiliar)								
Diseño del modelo (Dashboard indicador de violencia intrafamiliar)								
Evaluación indicador de violencia intrafamiliar (Correcciones por parte del director de pasantía)								
Puesta en producción del Dashboard de violencia intrafamiliar (publicación en la página web del observatorio)								

Fuente propia

Entendimiento de los datos

Proceso de análisis, filtración de datos para el diseño de los Dashboard

Cada uno de los diferentes documentos .xlsx que se presentaron a lo largo del proyecto, fueron otorgadas por distintas dependencias con las que tiene convenio el Observatorio socioeconómico de la Alcaldía de Fusagasugá, el proceso de realización del análisis del contenido, filtración, diseño de estadísticas, obtención de coordenadas, diseño de mapas y presentación por medio de Dashboard, son similares en los diferentes apartados de información, teniendo cada una de estas algunas particularidades propias a lo largo de su manipulación, pero todas con un fin común.

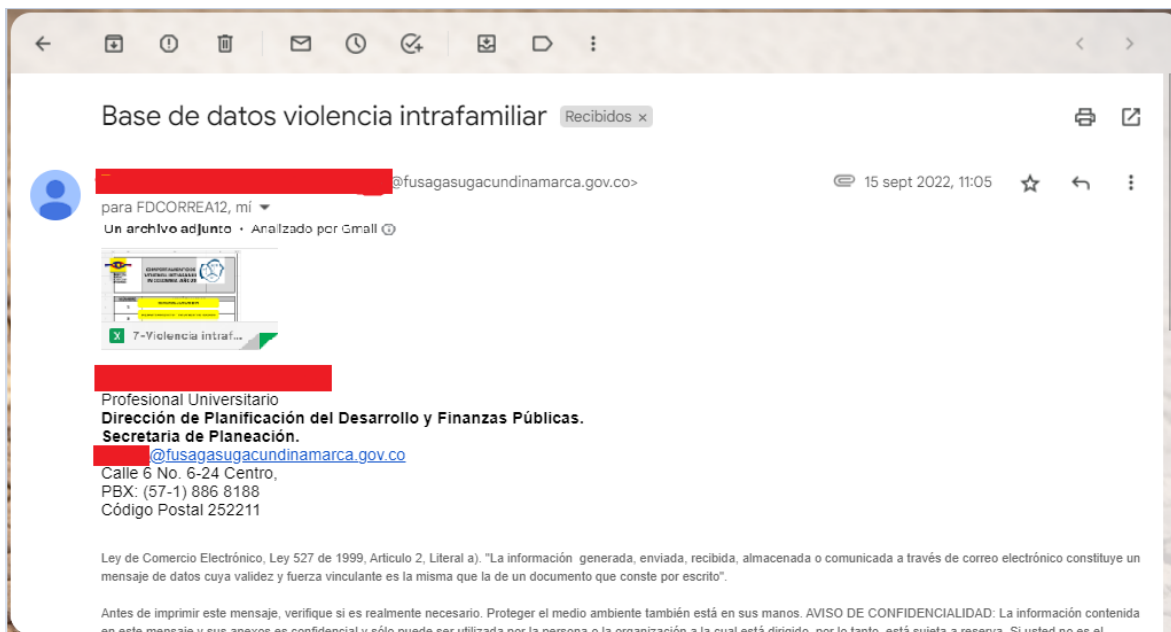
Información proporcionada por la organización (formatos y medio de envió)

Como ya se mencionó con anterioridad a lo largo de este documento, los registros de información fueron proporcionadas por diferentes dependencias de la alcaldía, cada una de estas fue proporcionada por medio de documentos de extensión .xlsx (Excel) albergando los datos correspondientes de las diferentes encuestas y estudios realizados por cada una de las oficinas que hacen parte de la alcaldía.

El medio por el cual fueron enviados los diferentes documentos fue por medio del correo electrónico oficial del observatorio socioeconómico de la alcaldía de Fusagasugá, como se ejemplifica en la siguiente ilustración:

Ilustración 9.

Medio de envío de documentos de almacenamiento de información de las dependencias.



Fuente: Propia

Proceso de análisis e interpretación de los datos almacenados

Cada uno de los datos a los que se tenía acceso contenía información única; para poder realizar una correcta gestión de la información era necesario tener contexto de la temática manejada, es por ello que, fue necesario hacer la interpretación de los datos, cada vez que se trabajaba con un documento nuevo, esto evitaba realizar operaciones innecesarias, erróneas o terminar omitiendo algún campo de información de gran importancia.

Al momento de haber la interpretación de los datos, ya era posible identificar ciertos datos que podían ser candidatos para el diseño de las estadísticas, se lograba determinar qué datos no eran candidatos para el desarrollo del producto, y ayudaba a generar una guía de cómo podría estructurarse el Dashboard.

Análisis de la información almacenada (Manejo de Google Colab)

La primera actividad realizada en el proceso de construcción de los Dashboard partió del diseño de los documentos de análisis de datos por medio de la herramienta de Google Colab y haciendo uso básico del lenguaje de programación Python. La elaboración de las actividades de análisis de datos consistía en la construcción de un documento para el estudio de la información con la que se disponía, con la finalidad de entender el contenido del indicador:

Importación del documento a Google Colab

Para la realización del documento de análisis de datos, se debe hacer una previa importación del documento desde Google Drive (Espacio donde se almacenan la información), para ejecutar esta tarea se ejecutan los siguientes bloques de código.

- Enlace a Google Drive

```
from google.colab import drive
drive.mount('/content/drive')
```

- Importación de librerías

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

```
import seaborn as sb
%matplotlib inline
```

- Conexión de Google Colab con el apartado donde se almacena el documento Excel con los datos.

```
data=pd.read_csv('/content/drive/MyDrive/ANALISIS ESTADISTICO/OBSERVATORIO/Secretaria de Movilidad/Accidentalidad_2019.csv', engine='python', encoding='ISO-8859-1', delimiter=';')
```

Complejidad de datos almacenados

Existe una línea de código, que al ser ejecutada arroja una lista con cada uno de los campos que se encuentran dentro de documento Excel, frente a estos campos se aprecia la cantidad de datos nulos que se encuentran, mostrando los registros que están incompletos.

Ilustración 10.

Código de datos nulos.

```
data.isnull().sum()
Fecha 0
Semestre 2019 0
Trimestral 2019 0
Longitud 0
Latitud 0
Codigo_provincia 0
Provincia 0
Cod_mun 0
Municipio 0
CARACTERIZACION POR CATEGORIAS 0
CARACTERIZACION DEL DELITO 0
GENERO 201
EDAD 392
dtype: int64
```

Fuente: Propia

Como se aprecia en la ilustración se puede observar que existe incompletitud en algunos campos del documento Excel (Genero y Edad).

La ausencia de estos campos genero cambios significativos con los resultados expuestos dentro de las estadísticas, ya que se presentó una ausencia de información en dos campos de gran importancia, como los son género y edad, ya que son datos candidatos para ser usados dentro los filtros, estadísticas o cantidades totales que se pueden exponer dentro de un Dashboard o dentro de un informe del Observatorio.

Consistencia de datos

Con la visualización de los datos por medio de tablas, con ayuda de una línea de código, se podía apreciar que cada uno de los campos almacenados dentro del documento se complementan uno con otro (Se llega a esta conclusión por medio de un análisis propio y un estudio de los datos), por lo que, la información almacenada no contiene contradicciones y cumple con este requerimiento de calidad.

Ilustración 11.

Tabla para análisis de consistencia de datos.

IPAT	FECHA DE ACCIEDNTE	DIRECCION	BARRIO	COMUNA	CORREGIMIENTO	GENERO	HORA	GRAVEDAD	
0	A0008244654	02/01/2019	CALLE 4 3 59	EMILIO SIERRA	CENTRO	NaN	MASCULINO	4:30:00 a. m.	SOLO DAÑOS
1	A001024278	03/01/2019	CARRERA 7 CON CALLE 10	CENTRO	CENTRO	NaN	MASCULINO	2:40:00 a. m.	HERIDOS
2	A001024280	04/01/2019	CALLE 25 48 50	COOVI PROF	SUR OCCIDENTAL	NaN	MASCULINO	11:40:00 a. m.	SOLO DAÑOS
3	A001024279	04/01/2019	CARRERA 3 22 17	FUSACATAN	SUR ORIENTAL	NaN	MASCULINO	1:20:00 a. m.	SOLO DAÑOS
4	A001024281	05/01/2019	TRANSVERSAL 12 ANTIGUA PANAMERICANA	NaN	SUR OCCIDENTAL	NaN	MASCULINO	2:45:00 a. m.	SOLO DAÑOS
...
284	A001050982	27/12/2019	CRA 64 22 05	MAIZ AMARILLO	SUR OCCIDENTAL	NaN	MASCULINO	4:47:00 p. m.	SOLO DAÑOS
285	A001050983	28/12/2019	CLLE 1A NORTE 9 18	LAS MARGARITAS	NORTE	NaN	MASCULINO	7:30:00 p. m.	HERIDOS

Fuente propia

Precisión/exactitud

Para verificar la precisión de los registros no se contó con una fuente de referencia, por lo que un medio para determinar este estándar de calidad, fue por medio de la comparación de registros, ya que dentro de los datos almacenados existieron registros que destacaban ser distintos a los otros, ya sea por falta de información o por tener un datos que no tenía relación con los demás, es por ello que la precisión y exactitud de los datos no estaba presente en todos los documentos que almacenaban la información.

Para ejemplificar la precisión/exactitud se pudo utilizar el mismo bloque de código usado para verificar la dimensión anterior (data).

Ilustración 12.

Tabla para análisis de exactitud de datos.

data					
	NÚMERO C	AREA DEFUN	SEXO FALLE	EDAD FALLE	MUNICIPIO
0	725742500	CABECERA MUNICIPAL	MASCULINO	60 o mas Años(Adulto mayor)	FUSAGASUGA
1	725742496	CENTRO POBLADO	MASCULINO	60 o mas Años(Adulto mayor)	FUSAGASUGA
2	725742482	CABECERA MUNICIPAL	MASCULINO	60 o mas Años(Adulto mayor)	FUSAGASUGA
3	725742518	CABECERA MUNICIPAL	FEMENINO	60 o mas Años(Adulto mayor)	FUSAGASUGA
4	725731888	CABECERA MUNICIPAL	MASCULINO	60 o mas Años(Adulto mayor)	FUSAGASUGA
...
1694	729601109	CABECERA MUNICIPAL	MASCULINO	60 o mas Años(Adulto mayor)	ARBELÁEZ
1695	729601130	CABECERA MUNICIPAL	FEMENINO	60 o mas Años(Adulto mayor)	PASCA
1696	729601205	CABECERA MUNICIPAL	FEMENINO	60 o mas Años(Adulto mayor)	SILVANIA
1697	729601283	CABECERA MUNICIPAL	INDETERMINADO	60 o mas Años(Adulto mayor)	BOGOTÁ, D.C.
1698	729601290	CABECERA MUNICIPAL	FEMENINO	60 o mas Años(Adulto mayor)	BOGOTÁ, D.C.

Fuente propia

Como se puede apreciar en la ilustración en el caso del campo de “Municipio”, se cuenta con una inconsistencia, ya que todos los datos con los que se deben trabajar deben ser del municipio de Fusagasugá, así que, todos los registros externos no deberían estar presentes en los almacenamientos de información de las dependencias, por lo que los registros no son exactos.

Conformidad de los datos

Para el almacenamiento de los datos no se contó con un formato estándar, lo que provocó que cada dependencia trabajara con un formato estructurado por ellos, esto interviene en los procesos de preparación de los datos, ya se tuvo que realizar una transformación de datos diferente para cada documento, como se puede apreciar con las siguientes ilustraciones.

Formato para almacenamiento de datos de nacimientos.

Ilustración 13.

Formato de almacenamiento de nacimientos.

1. DEFINICIÓN DEL INDICADOR				
CÓDIGO	NOMBRE DEL INDICADOR			
	PORCENTAJE DE NACIDOS VIVOS EN MUJERES ENTRE 14 Y 17 AÑOS.			
DESCRIPCIÓN INDICADOR		OBJETIVO DEL INDICADOR		
Número de nacidos vivos ocurridos en mujeres entre 14 y 17 años de edad.		Conocer el número de nacimientos ocurridos en mujeres entre 14 y 17 años.		
DIMENSION	SECTOR	DEPENDENCIA	OBJETIVO DE DESARROLLO SOSTENIBLE (ODS)	
DESARROLLO SOCIAL/CONVIVENCIA/CONSTR DE PAZ	SALUD	SECRETARÍA DE SALUD	SALUD Y BIENESTAR	
2. VARIABLES PARA MEDICIÓN DEL INDICADOR				
UNIDAD DE MEDIDA	LÍNEA BASE	PERIODICIDAD MEDICIÓN	FUENTE DE INFORMACIÓN	
Porcentaje	2010	Anual	Estadísticas vitales RUAF - Proyecciones de población DANE.	
FORMULA DE CALCULO				
Cociente entre el número de nacidos vivos ocurridos en mujeres entre 14 y 17 años y el total de nacimientos ocurridos en el año en el municipio de Fusagasugá.				
DESCRIPCIÓN NUMERADOR		DESCRIPCIÓN DENOMINADOR		
Total de niños y niñas nacidos de mujeres entre 14 y 17 años.		Total de niños y niñas nacidos en el municipio de Fusagasugá en un año.		
3. REPORTE DATOS				
AÑO	PERIODO	DATO NUMERADOR	DATO DENOMINADOR	RESULTADO MEDICIÓN
2010	ANUAL	35	352	9.94
2011	ANUAL	156	1332	11.71

Fuente propia

Formato para almacenamiento de datos de violencia intrafamiliar

Ilustración 14.

Formato de almacenamiento de datos de violencia intrafamiliar.

DIVIPOLA	VIOLENCIA INTRAFAMILIAR (VIF) EN COLOMBIA AÑO 2020 NÚMERO DE CASOS	VIOLENCIA CONTRA NIÑOS, NIÑAS Y ADOLESCENTES (VIF)						TOTAL
		Violencia contra Niños, Niñas y Adolescentes (VIF)	Violencia contra Niños, Niñas y Adolescentes (VIF)	Violencia contra Niños, Niñas y Adolescentes (VIF)	Violencia contra Niños, Niñas y Adolescentes (VIF)	Violencia contra Niños, Niñas y Adolescentes (VIF)	Violencia contra Niños, Niñas y Adolescentes (VIF)	
CÓDIGOS	← INICIO SEXO	Hombre	Hombre	HOMBRE	Mujer	Mujer	MUIER	TOTAL
N°	DEPARTAMENTO Y MUNICIPIO DEL HECHO DANE / MAYOR DE EDAD (> 18 AÑOS) O MENOR DE EDAD (< 18 AÑOS)	Menores de Edad (<18 años)	Mayores de Edad (>18 años)	TOTAL	Menores de Edad (<18 años)	Mayores de Edad (>18 años)	TOTAL	
25	Cundinamarca	225	-	225	229	-	229	454
25290	Fusagasugá	10	-	10	11	-	11	21
0	Colombia	2,392	-	2,392	2,677	-	2,677	5,069

Fuente propia

Duplicación de datos

Dentro de los registros no se debe presentar la repetición de la información (dependiendo del dato), ya que se puede presentar irregularidades dentro del Dashboard que se estaba construyendo.

Como se puede apreciar en la siguiente ilustración, se puede apreciar la duplicación de datos, hasta tres registros iguales, esto pudo provocar que la información publicada contuviera datos erróneos, es por esto que, al momento de hacer la preparación de los datos se debió verificar los registros para determinar su dimensión de calidad de duplicación.

Ilustración 15.
Duplicación de datos.

Nombre del establecimiento comercial:	Dirección del establecimiento:	Latitud
Quesos somos colombia	Plaza de mercado 2 piso local 199	4.342055
La Dora restaurante campestre	Km 4 vida jordan vía la Aguadita	4.388256
Soda Burger	Cra 11 a # 16 b 34	4.336585
Farmacenter maiz amarillo	Cra 64. 23b40 local 101	4.329555
Farmacenter maiz amarillo	Cra 64. 23b40 local 101	4.329555
El maná express	Carrera 2 # 6-44	4.343919
El maná express	Carrera 2 # 6-44	4.343919
El maná express	Carrera 2 # 6-44	4.343919
Expendio de carnes CB	Carrera 12 # 7 a 20	4.342607
Farmacenter coviprof	Cra 39b. 24-12 local 101	4.341054
Merka Exito	Calle 23#39-08	4.328353
El Antojo	Calle 17 A #12-14 Piedra Grande, Fusagasuga	4.335979
Farmacenter villaceleste	Cra 75. 24-05	4.666161
Panadería la excelencia	Calle 17A#12-54	4.335837
Farmacenter Terranova	Cra 64. 22 d 21	4.33301
Rogan Futuro sas	Fusagasuga	4.336736
Pollo saludable	Vda jordan via la aguadita	4.363799
Efecty - coresponsal Bancolombia -corr	Calle 20 # 4 78	4.329481
Plaza de mercado	Plaza de mercado	4.342003

Fuente propia.

Con ayuda del análisis de datos se pudo determinar la calidad con la que contaban los datos que estaban compartiendo las dependencias de la alcaldía, resaltando diferentes inconformidades de los registros, comparándolos con las dimensiones de calidad que debe tener los datos. Además de esto, gracias a los análisis que se realizaron, se aportó información para la construcción de los diferentes informes que publica el observatorio socioeconómico de la alcaldía de Fusagasugá dentro de su página oficial <https://observatorio.alcaldiafusagasuga.gov.co/informes-2/>, (La construcción de los informes es realizada por el encargado del área estadística del observatorio no por el pasante).

Preparación de los datos

Selección, conversión y corrección de datos para el Dashboard

Antes de poder iniciar con el diseño las estadísticas, los Dashboard y las historias (Conjunto de Dashboard almacenados en un mismo segmento), Era necesario realizar una serie de correcciones a los documentos que contenían la información.

Ya que, para poder implementar de manera correcta la información dentro del programa (Tableau), se debía realizar un proceso de extracción y transformación de datos, además de, Suprimir ciertas columnas innecesarias dentro del documento de extensión Excel, para entender mejor el proceso de transformación de los medios de almacenamiento, se mostrará un documento que almacena información proporcionada por la secretaria de salud de Fusagasugá.

Ilustración 16.

Documento sin corrección.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1		2020														
2	VIOLENCIA INTRAFAMILIAR (VIF) EN COLOMBIA AÑO 2020 / MUNICIPIO DEL HECHO DANE	Cundinamarca / Fusagasugá	Cundinamarca / Fusagasugá	Cundinamarca / Fusagasugá	Cundinamarca / Fusagasugá	Cundinamarca / Fusagasugá	Cundinamarca / Fusagasugá	Cundinamarca / Fusagasugá								
3	CÓDIGOS DIVIPOLA	###	###	###	###	###	###	###								
4	← INICIO SEXO	Homere	Homere	HOMBRE	Mier	Mier	MUER	TOTAL								
5	← INICIO VIOLENCIA - VARIABLE / MAYOR DE EDAD (> 18 AÑOS) O MENOR DE EDAD (< 18 AÑOS)	Mayor de 18 años	Mayor de 18 años	TOTAL	Mayor de 18 años	Mayor de 18 años	TOTAL									
563	VARIABLE: RANGO DE HORA DEL HECHO (X 3 HORAS)															
564	VIOLENCIA CONTRA NIÑOS, NIÑAS Y ADOLESCENTES (VIF)	10	-	10	11	-	11	21								
565	(00:00 a 02:59)	-	-	-	-	-	-	-								
566	(03:00 a 05:59)	-	-	-	-	-	-	-								
567	(06:00 a 08:59)	-	-	-	2	-	2	2								
568	(09:00 a 11:59)	2	-	2	3	-	3	5								
569	(12:00 a 14:59)	3	-	3	1	-	1	4								
570	(15:00 a 17:59)	2	-	2	1	-	1	3								
571	(18:00 a 20:59)	3	-	3	2	-	2	5								
	←	MENÚ VIF 2020	DEPARTAMENTO Y MUNICIPIO CASOS	DEPARTAMENTO Y MUNICIPIO TASAS	MUNICIPIO DEL HECHO											

Fuente: Propia.

En la ilustración del documento que almacena información proporcionada por la secretaria de salud se puede encontrar los siguientes componentes innecesarios para el proceso de construcción de los Dashboard:

- Uso de símbolos matemáticos para resaltar los datos nulos.
- Uso de columnas innecesarias.
- Uso de títulos innecesarios para describir los datos.
- Información de poca importancia social.
- Columnas de datos repetidas

La corrección de la forma en la que se organiza la organización da como resultado un documento con:

- Exclusión de símbolos matemáticos innecesarios
- Almacenamiento de datos de gran interés ya sea para usuarios o para la realización de un estudio por parte de alguna dependencia (filtración de datos)
- Implementación de columnas nuevas para almacenar datos que tiene relación entre si (Transformación).
- Omisión de datos nulos.
- Implementa el uso de una sola columna para presentar un tipo de datos (en el anterior documento se presentaban varias columnas para un solo tipo de dato).

El resultado de las correcciones las podemos apreciar en la siguiente ilustración.

Ilustración 17.
documento corregido.

Día del hecho	Categoría Día del hecho	Cantidad Día del hecho	Genero Día del hecho	Codigo Día del hecho
lunes	Niños, niñas y adolescentes	1	Masculino	1
martes	Niños, niñas y adolescentes	2	Masculino	1
miércoles	Niños, niñas y adolescentes	2	Masculino	1
jueves	Niños, niñas y adolescentes	1	Masculino	1
sábado	Niños, niñas y adolescentes	3	Masculino	1
domingo	Niños, niñas y adolescentes	1	Masculino	1
lunes	Niños, niñas y adolescentes	3	Femenino	1
martes	Niños, niñas y adolescentes	1	Femenino	1
miércoles	Niños, niñas y adolescentes	3	Femenino	1
jueves	Niños, niñas y adolescentes	1	Femenino	1
viernes	Niños, niñas y adolescentes	1	Femenino	1
sábado	Niños, niñas y adolescentes	1	Femenino	1
domingo	Niños, niñas y adolescentes	1	Femenino	1
lunes	otros familiares	2	Masculino	1

Fuente: Propia.

Selección de datos

La primera actividad que se va a realizar para poder construir el Dashboard, es la selección de los datos, Como se especificó en los procesos anteriores, es necesario contar con el contexto de la información con la que estaba trabajando, para poder determinar qué datos usar para la construcción de las estadísticas. Para dar una mejor explicación de cómo fue realizado la selección de datos en los documentos que almacenaban la información, se presentará una tabla que describe los campos que se encontraban en un documento Excel compartido por una de las dependencias de la alcaldía (secretaria de comercio), este documento fue proporcionado por la secretaria de comercio (establecimientos comerciales).

Dato	Tipo de dato	Descripción
Marca temporal	DATE	Describe la fecha de realización de la encuesta, exponiendo el día, el mes, el año y la hora de realización
Tipo de establecimiento comercial que posee o administra	VARCHAR	Expone el tipo de establecimiento tomando como referencia una serie de categorías (Droguería, Fruver, Panadería, etc.)
Nombre del establecimiento comercial	VARCHAR	Nombre del establecimiento comercial
Nombres y apellidos del propietario o administrador del establecimiento:	VARCHAR	Nombres y apellidos del dueño del establecimiento comercial
Teléfono de contacto:	INT	Teléfono de contacto personal del dueño del establecimiento comercial
Teléfono de contacto, para que sus clientes lo contacten	INT	Teléfono de contacto del establecimiento comercial
Dirección del establecimiento:	VARCHAR	Dirección de ubicación del establecimiento comercial
Latitud	INT	Latitud de la ubicación del establecimiento comercial (Este dato se implementó por medio de un proceso de conversión)
Longitud	INT	Longitud de la ubicación del establecimiento comercial (Este dato se implementó por medio de un proceso de conversión)
Barrio:	VARCHAR	Barrio de la ciudad de Fusagasugá en el que se encuentra ubicado el establecimiento
Productos que comercializa en su establecimiento	VARCHAR	Tipos de productos que se comercializan en el establecimiento (dentro de este dato se tuvo que hacer una división de datos para poder implementarse dentro de Tableau)
Realiza domicilios:	VARCHAR	Establecimientos que realizan o no domicilios (se divide en sí, No, pero sí los realizaría, en caso de ser necesario y No, y no los realizaría)

Tabla 13

Dentro de esta lista de opciones de datos para escoger, se tuvo en cuenta los datos de información de identificación personal, al encontrarnos con datos de este tipo, se hizo un descarte inmediato, ya que este tipo de datos no son candidatos para publicar en la página de la alcaldía, ya que podría conllevar a problemas de índole legales.

Dentro de este documento se encontraron los siguientes datos de tipo PII.

- Nombres y apellidos del propietario o administrador del establecimiento
- Teléfono de contacto
- Nombre del establecimiento comercial
- Teléfono de contacto, para que sus clientes lo contacten

Conociendo estos datos, se pudo continuar con la selección de datos, y se determinó que los datos óptimos para la construcción del Dashboard fueron:

- Recuento de campos “200” (Este campo se integró dentro de la aplicación Tableau).
- “Realiza domicilios:”
- Latitud
- Longitud
- Barrio:
- Productos que comercializa en su establecimiento:

Como se aprecia dentro de la Tabla 11 (Campos del documento “Establecimientos comerciales”), existen dos campos de localización geográfica (latitud y longitud), estos datos fueron implementados realizando un proceso de conversión de direcciones que se explicara a continuación.

Proceso para obtener la latitud y longitud con base a las direcciones

Dentro de los procesos de preparación de datos, fue necesario realizar una transformación de los datos de direcciones, para obtener datos de tipo latitud y longitud, ya que, para la construcción de los diferentes mapas es necesario tener los datos en este medio.

En un inicio se planteó realizar la transformación de los datos por medio del uso de Google Colab con ayuda de la biblioteca de geopy, pero este medio contaba con una restricción de cantidad de datos, lo que causaba que, realizar la transformación de una gran cantidad de datos fuera una actividad muy demorada, la conversión de treinta datos podía tomar hasta treinta minutos. Es por ello que, la gestión de documentos de más de mil datos sería una actividad muy demorada.

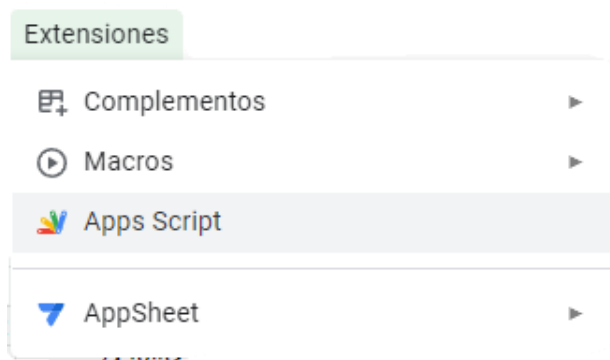
Para realizar la tarea de conversión de datos se propuso hacer uso de las hojas de cálculo de Google (Google Sheets), dado que, dentro de esta herramienta se puede hacer el uso de codificación y librerías para ejercer diferentes actividades, por medio del editor de secuencias de mandos (Script Editor).

La configuración de Script Editor fue realizada por el autor del video y pagina web "Tesel", propuso un método de transformación de datos en formato de direcciones, para poder obtener la latitud y longitud de dichas ubicaciones, la codificación se dividió en configuración del proyecto, geo codificación con ayuda de Google maps e impresión de resultados. (Tesel - 2021).

El primer paso por realizar para obtener la latitud y longitud es, acceder al apartado de App Script, en el campo de extensiones de la barra de tareas de Google Sheets.

Ilustración 18.

Barra de extensiones de Google Sheets

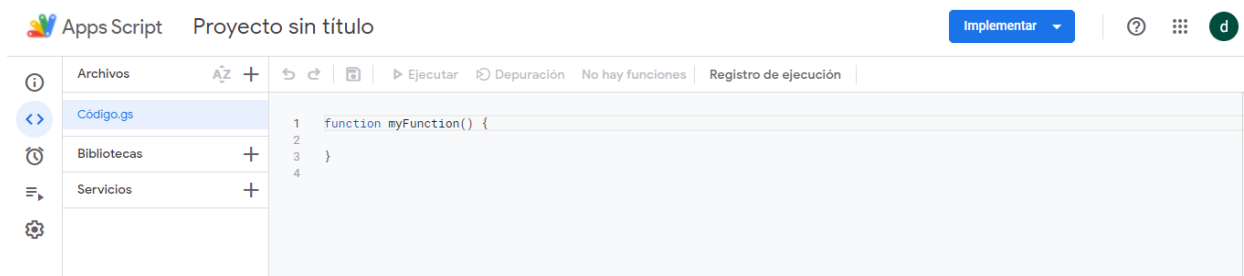


Tomado de: Google Sheets. Apartado de extensiones de la barra de tareas.

Al acceder a la pestaña de App Script se realizará una redirección a una pestaña con un apartado para donde se realizó la programación, por medio del uso de lenguaje Java Script (con características y funciones especiales).

Ilustración 19.

Interfaz de App Script



Tomado de: Google Sheets. App Script.

Menú para la geo codificación

Para realizar la ejecución del código e iniciar la transformación de las direcciones, primero se tiene que crear el ítem que iniciara todo el proceso, las siguientes líneas de código realizan esta función antes mencionada, agregando un botón con el nombre que se desee (en este caso Transformar) en la barra de tareas de Google Sheets.

```
var ss = SpreadsheetApp.getActiveSpreadsheet();
var sheet = ss.getActiveSheet();
var dataRangeAll = sheet.getDataRange();
var ultimaFila = dataRangeAll.getLastRow();

function onOpen() {
  var ui = SpreadsheetApp.getUi();
  ui.createMenu('Transformar')
    .addItem('Tranformar las direcciones en coordenadas (latitud y longitud)', 'Transfo
rmar')
    .addToUi();
}
```

Transformación de las direcciones

Ahora para convertir los datos se realiza hacer un recorrido de todas las filas, partiendo de la segunda, ya que la primera es el título, con ayuda de la herramienta Google maps, para determinar la latitud y longitud de los textos ingresados en la primera columna de la hoja de cálculo, y obtener los resultados en la segunda y tercera columna del documento, además de esto, dentro del código se implementó un pequeño espacio para realizar la ejecución solo en los campos con información, excluyendo campos nulos para evitar fallos en la ejecución.

```
function Transformar() {
  var filaInicial = 2;
  var dataRange = sheet.getRange(filaInicial, 1,ultimaFila, 3);
  var data = dataRange.getValues();

  for(var i=0; i< data.length; i++) {
    var fila = data[i];
    var direccion = fila[0];
    if(direccion != "") {
```

```

// Conexión con Google Maps
var geocoder = Maps.newGeocoder().geocode(direccion);
var resultado = geocoder.results[0];

var latitud = 0;
var longitud = 0;

if(resultado) {
    latitud = resultado.geometry.location.lat;
    longitud = resultado.geometry.location.lng;
    sheet.getRange(filaInicial + i, 2).setValue(latitud);
    sheet.getRange(filaInicial + i, 3).setValue(longitud);
}
}
}
}

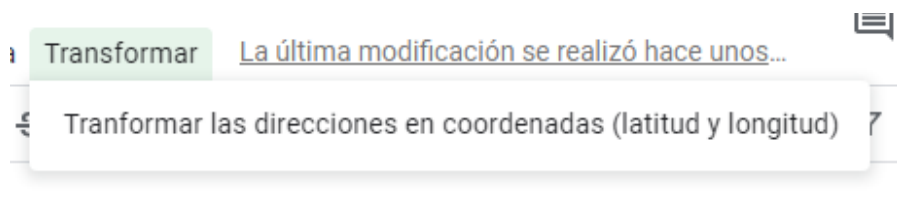
```

Una vez ejecutado el bloque de código, dentro de la hoja de cálculo de Google Sheets aparecerá un nuevo botón en la barra de tareas, denominado Transformar, es aquí donde se inicia la ejecución del código realizado en App Script.

Nota: se debe ingreso las direcciones en la primera columna, pero para obtener un dato más exacto se agregó el país y ciudad de donde provenían las direcciones, separando los datos por medio de comas para que no allá errores en la ejecución.

Ilustración 20.

Botón de transformación de datos.



Tomado de: Documento de Google Sheets

Fuente propia.

Ilustración 21.

Direcciones para obtención de coordenadas latitud y longitud.

Dirección del establecimiento:
Colombia,Fusagasugá,Calle 5 # 9 - 06
Colombia,Fusagasugá,Calle 22 #8-02
Colombia,Fusagasugá,calle 21# 22-41
Colombia,Fusagasugá,Carrera 5 #10-15
Colombia,Fusagasugá,Carrera 1 # 3-16
Colombia,Fusagasugá,carrera 11 a # 16 b 34
Colombia,Fusagasugá,carrera64. 23b40 local 101
Colombia,Fusagasugá,carrera 64. 23b40 local 101
Colombia,Fusagasugá,Carrera 2 # 6-44
Colombia,Fusagasugá,Carrera 2 # 6-44
Colombia,Fusagasugá,Carrera 2 # 6-44
Colombia,Fusagasugá,Carrera 12 # 7 a 20
Colombia,Fusagasugá,carrera 39b. 24-12 local 101
Colombia,Fusagasugá,Calle 23#39-08
Colombia,Fusagasugá,Calle 17 A #12-14
Colombia,Fusagasugá,carrera 75. 24-05
Colombia,Fusagasugá,Calle 17A#12-54
Colombia,Fusagasugá,Cra 64. 22 d 21
Colombia,Fusagasugá,Calle 20 # 4 78

Tomado de: Documento de Google Sheets

Fuente propia.

Al momento de ejecutar el código desarrollado en App Script se arrojará los datos latitud (columna dos) y longitud (columna tres), y gracias a esto se podrá realizar la construcción de los mapas de georreferenciación de información para el desarrollo de los Dashboard, para publicar dentro de la página oficial del observatorio socioeconómico de la alcaldía de Fusagasugá.

Ilustración 22.

Direcciones para obtención latitud y longitud.

Dirección del establecimiento:	Latitud	Longitud
Colombia,Fusagasugá,Calle 5 # 9 - 06	4.347025	-74.357706
Colombia,Fusagasugá,Calle 22 #8-02	4.342055	-74.364442
Colombia,Fusagasugá,calle 21# 22-41	4.388256	-74.325016
Colombia,Fusagasugá,Carrera 5 #10-15	4.336585	-74.36554
Colombia,Fusagasugá,Carrera 1 # 3-16	4.329555	-74.399109
Colombia,Fusagasugá,carrera 11 a # 16 b 34	4.329555	-74.399109
Colombia,Fusagasugá,carrera64. 23b40 local 101	4.343919	-74.358555
Colombia,Fusagasugá,carrera 64. 23b40 local 101	4.343919	-74.358555
Colombia,Fusagasugá,Carrera 2 # 6-44	4.343919	-74.358555
Colombia,Fusagasugá,Carrera 2 # 6-44	4.342607	-74.365619
Colombia,Fusagasugá,Carrera 2 # 6-44	4.341054	-74.360883
Colombia,Fusagasugá,Carrera 12 # 7 a 20	4.328353	-74.384728
Colombia,Fusagasugá,carrera 39b. 24-12 local 101	4.335979	-74.367837
Colombia,Fusagasugá,Calle 23#39-08	4.666161	-74.119755
Colombia,Fusagasugá,Calle 17 A #12-14	4.335837	-74.367698
Colombia,Fusagasugá,carrera 75. 24-05	4.33301	-74.397676
Colombia,Fusagasugá,Calle 17A#12-54	4.336736	-74.374202
Colombia,Fusagasugá,Cra 64. 22 d 21	4.363799	-74.359184
Colombia,Fusagasugá,Calle 20 # 4 78	4.329481	-74.364704

Tomado de: Documento de Google Sheets

Fuente propia.

Diseño de las estadísticas en Tableau

Después de haber realizado los procedimientos de los pasos anteriores, se inició con el diseño de las estadísticas para la construcción del Dashboard. en Tableau dentro de estas estadísticas se integraron los datos seleccionados en los precisos anteriores, al momento de hacer el grafico estadístico, se buscaba realizar un gráfico con información entendible y de gran interés para los usuarios.

Dentro del programa usado para el diseño de las estadísticas (Tableau) se tenían diferentes opciones para la elección del diseño (basados en la estructura de los datos), dentro de todos los Dashboard que se diseñaron se puede encuentran los siguientes tipos de estadísticas:

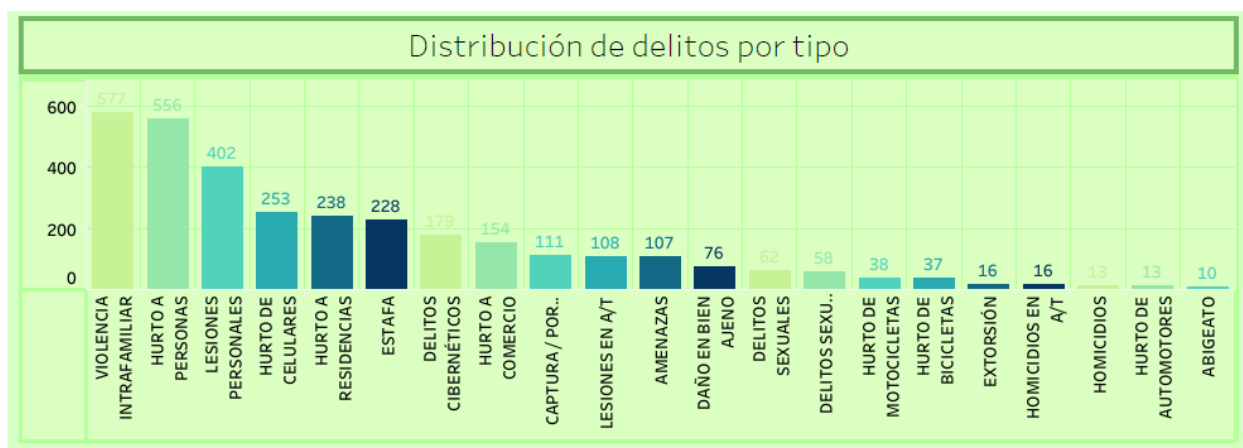
Barras verticales

Este tipo de graficas reflejaron los datos por medio de una serie de barras verticales de proporciones asociadas a la cantidad del dato, la siguiente ilustración muestra una distribución de delitos categorizados por tipo.

Nota: Esta información proviene del documento de “seguridad y convivencia del año 2020”.

Ilustración 23.

Diagrama de barras verticales.



Fuente: Propia.

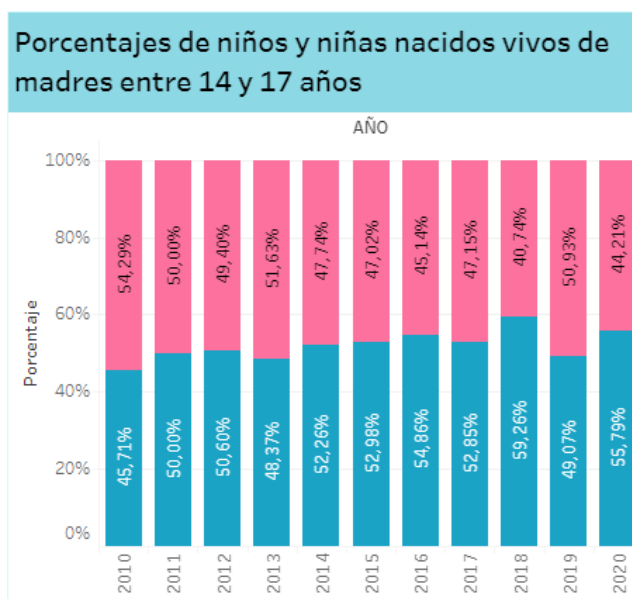
Gráfico de barras apiladas

Este tipo de graficas son un conjunto de barras rectangulares apiladas una sobre otra, las cuales dividen la información por medio de patrones de colores dependiendo de los datos. La siguiente Ilustración expone una gráfica aplicada,

La grafica muestra el porcentaje de niños nacidos vivos de madres entre 14 y 17 años en la ciudad de Fusagasugá dependiendo del año (Uso de varios datos para una solo grafica). La fuente de la ilustración proviene del documento que almacena información de “INDICADORES DE NACIMIENTOS”.

Ilustración 24.

Gráfico de barras apiladas.



Fuente: Propia

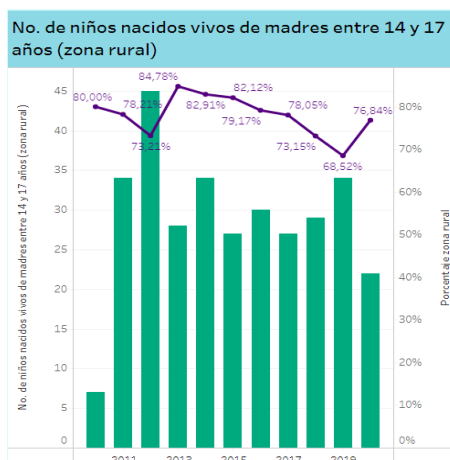
Combinación doble

Este tipo de grafica son una combinación de dos estilos de graficas para generar una sola, en el caso de este ítem se combinó de una gráfica de barras verticales y una gráfica de puntos.

Para ejemplificar este tipo de grafica se tomó el indicador de “INDICADORES DE NACIMIENTOS”, en la cual se muestra la cantidad de nacimientos, el año en que se tomó la información y por último el porcentaje de la zona rural.

Ilustración 25.

Gráfico de barras apiladas.



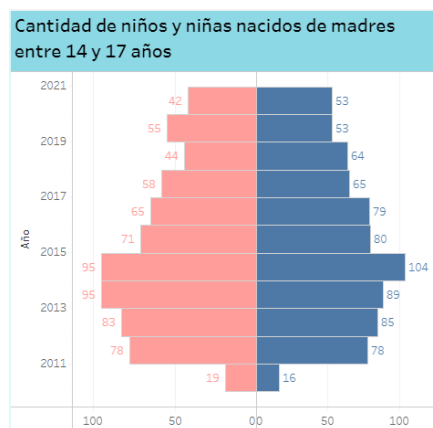
Fuente: Propia

Histograma

Es una gráfica que representa datos por medio de columnas proporcionales a la información, en la ilustración se muestra una gráfica de columnas horizontales dobles, que expone la información sobre la cantidad de niños nacidos de madres entre 14 y 17 años.

Ilustración 26

Gráfico de histograma



Fuente: Propia

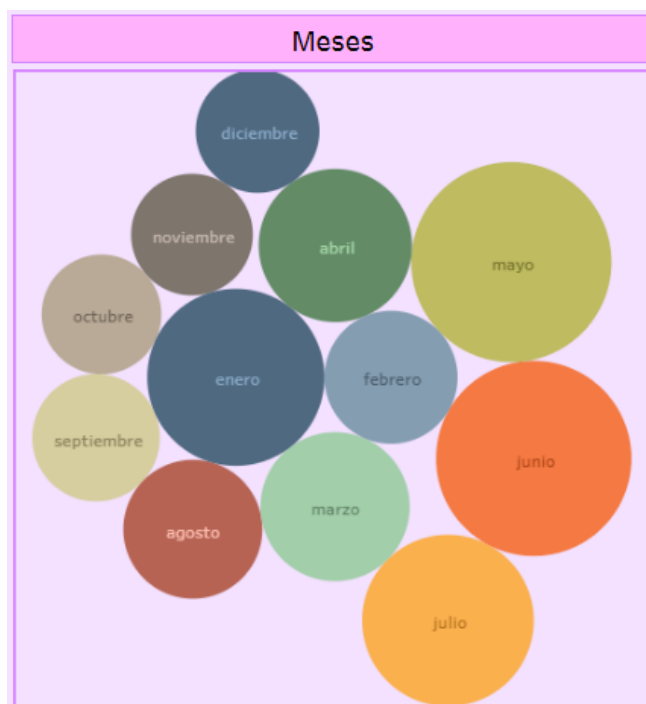
Grafica de burbujas agrupadas

Es una variante de un gráfico de dispersión, en el que los datos se presentan por medio de burbujas, de diferentes tamaños dependiendo de los datos que se están manejando como base.

Para ejemplificar este tipo de grafica se tomó como referencia el documento que almacena información de “Defunciones del año 2021”, donde se aprecia los diferentes círculos cada uno representando un mes del año y con diferentes tamaños dependiendo de la cantidad de fallecimientos ocurridos por medio de las burbujas, y el tamaño será definido por la cantidad de defunciones ocurridas.

Ilustración 27

Gráfico de burbujas agrupadas



Fuente: Propia

Gráfico de anillos

Es una variante del gráfico circular, en donde se presentan los datos en una figura circular con un espacio en blanco en el centro, asimilando la forma de una dona, dentro de este espacio en blanco se puede encontrar la exposición de otro dato ya sea por medio de palabras o de manera numérica.

Esta estadística divide los datos por medio de divisiones de color dependiendo del tamaño de los datos. Dentro de la ilustración se parecía la estadística resultante de la gestión de información del documento que almacena información de establecimientos comerciales, en el apartado de domicilios (Proporcionada por la secretaria de comercio), se puede apreciar de mejor manera el tipo de grafica.

Ilustración 28

Gráfico de burbujas agrupadas



Fuente: Propia

Apartados desarrollados con las herramientas disponibles dentro de Tableau

Dentro de Tableau existe una serie de herramientas que puede expandir de gran manera los diseños de los tableros, mencionando algunas de estas herramientas, tanto las del apartado gráfico, como las de desarrollo por medio de códigos se hizo uso de las siguientes herramientas.

Cantidades

Cuando mencionamos cantidades se hace referencia al conteo ya sea de campos, como de datos individuales por medio de la búsqueda de palabras concretas.

Para el desarrollo de los Dashboard se hizo uso de las dos opciones de conteo de cantidades, para ejemplificar el uso de las cantidades se mostrará el conteo de datos de género que se realizó en varios Dashboard (Masculino y Femenino), por medio del siguiente código.

Ilustración 29

Codificación para cantidades.



```

COUNT
(IF CONTAINS([Genero], 'MASCULINO') == TRUE THEN [Genero]
END)

```

El cálculo es válido. 5 dependencias ▼ Aplicar Aceptar

Fuente: Propia

El resultado que se arroja con el conteo de datos se puede apreciar en la siguiente ilustración, en donde se muestra el conteo de tres datos presentes en el documento Excel (cantidad, total hombres y total mujeres).

Ilustración 30

Herramienta de conteo para cantidades.



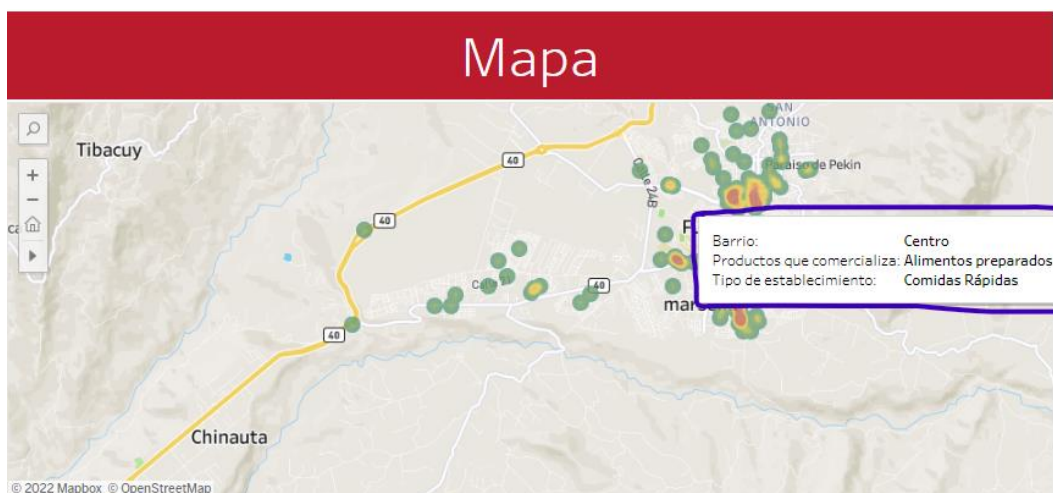
Fuente: Propia

Etiquetas

una etiqueta es un resumen de la información que se encuentra dentro de la estadística, y se hace visible si se coloca el cursor sobre en campo, columna o apartado de información de la estadística, o dentro del algún apartado o área del mapa de georreferenciación de información como se muestra a continuación.

Ilustración 31

Herramienta de etiquetas.



Fuente: Propia.

Desarrollo de mapas de georreferenciación de la información

Dentro de los diferentes componentes que se encuentran dentro de los Dashboard se aprecian una serie de mapas de georreferenciación de información, estos mapas se lograron implementar gracias a la transformación de los campos de las direcciones de los documentos .xlsx (Excel) que se estaba trabajando en ese momento, gracias a la herramienta de lectura de longitud y latitud con la que cuenta Tableau se logró estructurar una serie de mapas para los Dashboard.

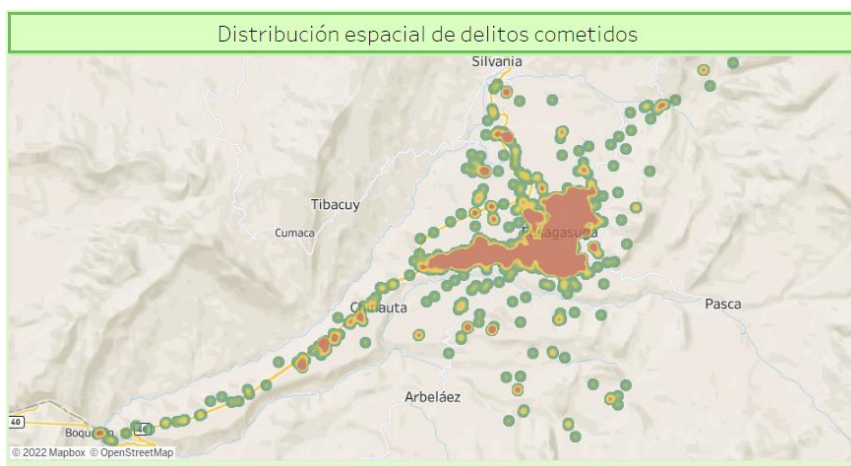
Algunos de los tipos de mapas que se usaron para el diseño de los Dashboard son:

Mapa de densidad

Es un mapa de calor que refleja una densidad de datos agrupados en un punto específico, cuando era la cantidad de datos agrupados en un mismo punto, el mapa iba tomando una paleta de colores rojizos, y si la agrupación era menor, se los tornara un color verde, indicando una menor agrupación de datos.

Ilustración 32

Mapa de calor (densidad).



Fuente: Propia.

Mapa de zona rural y urbana

Es un mapa que divide las zonas rurales y urbanas de la ciudad de Fusagasugá, dividiéndose por medio del uso de colores para la identificación de cada zona, como se puede apreciar en la Ilustración, la implementación de este tipo de mapa se realizó en la construcción de Dashboard para “Nacimientos”.

Concluyendo en mapa que divide la zona rural y la urbana por medio de dos colores, un color verde para identificar la zona rural y un color rosado para identificar la zona urbana.

Ilustración 33

Mapa de zona rural y urbana



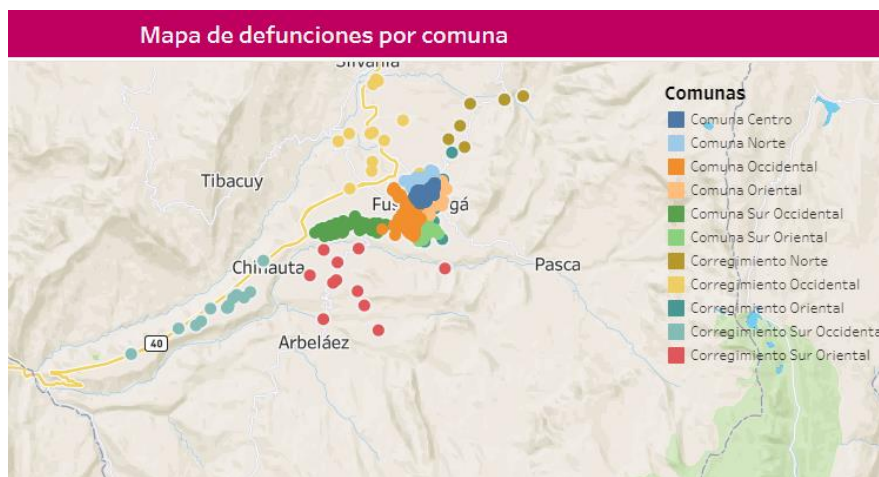
Fuente: Propia.

Mapa de comunas

Es un mapa dividido por comunas, especifica el sector de la comuna por medio de la separación con paleas de colores distintas para cada una de las comunas de la ciudad.

Ilustración 34

Mapa de división por comunas



Fuente: Propia.

Implementación de filtros de información

Dentro de los componentes de los Dashboard se encuentran los filtros, estos constan de una lista de datos para seleccionar, su función es modificar los datos mostrados en las estadísticas dependiendo de que valores escoja un usuario.

Un ejemplo de filtros se puede apreciar en el Dashboard de “defunciones del año 2021”, las opciones de filtración que se tienen se aprecian en la siguiente ilustración:

Ilustración 35

Filtro para Dashboard defunciones 2021.

Filtros

Rango de edad

- (Todo)
- 0-5 Años(Primera i...
- 6-11 Años(Infanci...
- 12-17 Años(Adol...
- 18-28 Años(Juve...
- 29-44 Años(Adult...
- 45-59 Años(Vejez)
- 60 o mas Años(A...

Genero

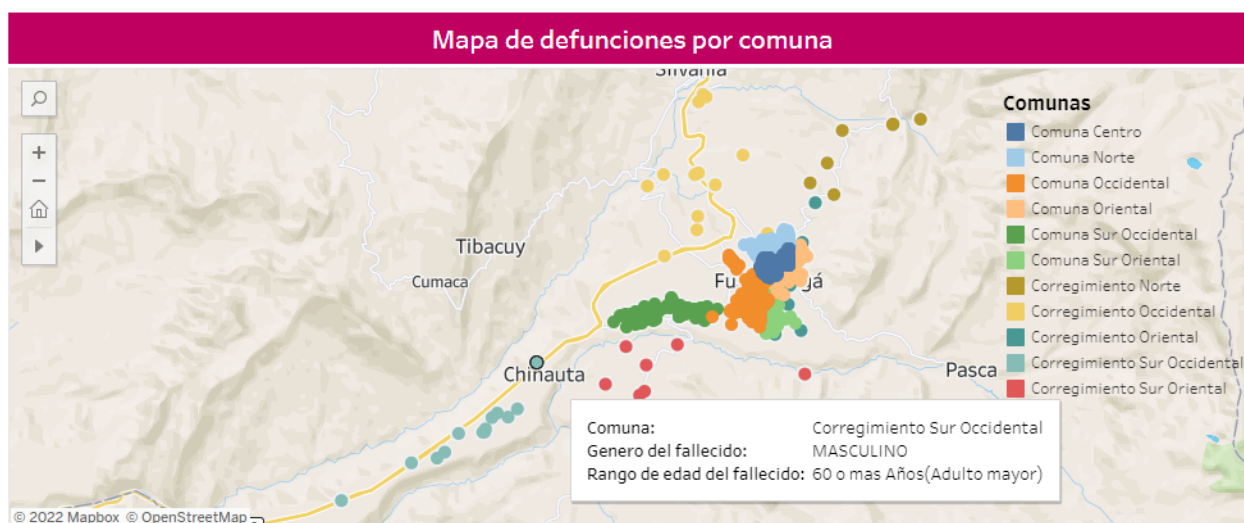
- (Todo)
- FEMENINO
- MASCULINO

Fuente: Propia.

Sin realizar la filtración de datos podemos observar dentro del Dashboard el siguiente mapa de comunas:

Ilustración 36

Mapa sin uso de filtro



Fuente: Propia.

Pero una vez que se realiza la filtración de datos por ejemplo con la selección de solo datos de personas mayores de 0 a 5 años podemos apreciar el siguiente cambio en el mapa:

Ilustración 37

Mapa con uso de filtro.



Fuente: Propia.

Creación de los modelos

Creación del Dashboard

Al momento en el que se diseñó el tablero o Dashboard, fue necesario contar con los objetos a implementar, esto quiere decir que, se debió tener a disposición las gráficas, mapas y demás apartados correctamente estructurados, diseñados y configurados para la implementación.

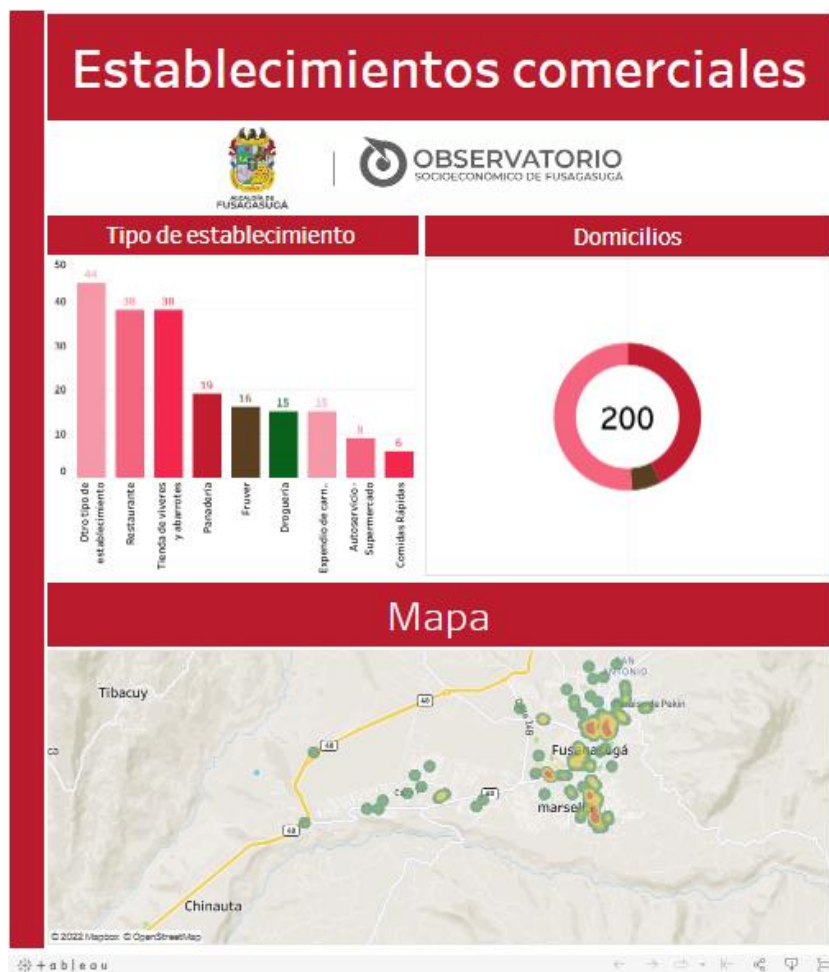
El Dashboard es el encargado de organizar y exponer la información, de una manera comprensible para los usuarios que lo verán, a lo largo de la realización se logró estructurar un gran número de estos Dashboard, cada uno con cantidades y tamaños diferentes en lo que respecta a datos e información, pero todos cumpliendo un propósito único, el cual es mostrar la información de manera organizada y comprensible.

Para realizar esta tarea se tuvo a disposición el apartado de creación de tableros de Tableau, en donde se incorporaron las estadísticas y mapas realizados con anterioridad, además de, una serie de ítems (etiquetas, filtros y conteo de datos), para que el diseño sea más dinámico e interactivo, se contó con un gran número de opciones para la configuración para los Dashboard desde la opción de implementación de títulos y texto, hasta la posibilidad de insertar imágenes.

Se realizó los diseños de cada producto pensando en una manera de poder organizar los datos, es por eso por lo que cada uno de estos tableros contaron con un diseño diferente, buscando tener un orden específico para que se acople mejor al tipo de información expuesta y también para que sea intuitivo su manejo y análisis por parte de las dependencias de la alcaldía y los usuarios comunes.

Un ejemplo de una Dashboard completamente terminado se puede apreciar en la siguiente ilustración, allí se evidencia la implementación de todas las estadísticas construidas en base a los procesos antes mencionados, exponiendo un mapa de georreferenciación de información y contando con un ítem de conteo de datos para que los usuarios sepan las cantidades totales de los datos.

Ilustración 38
Dashboard de establecimientos comerciales



Fuente: Propia.

Creación de historias en Tableau

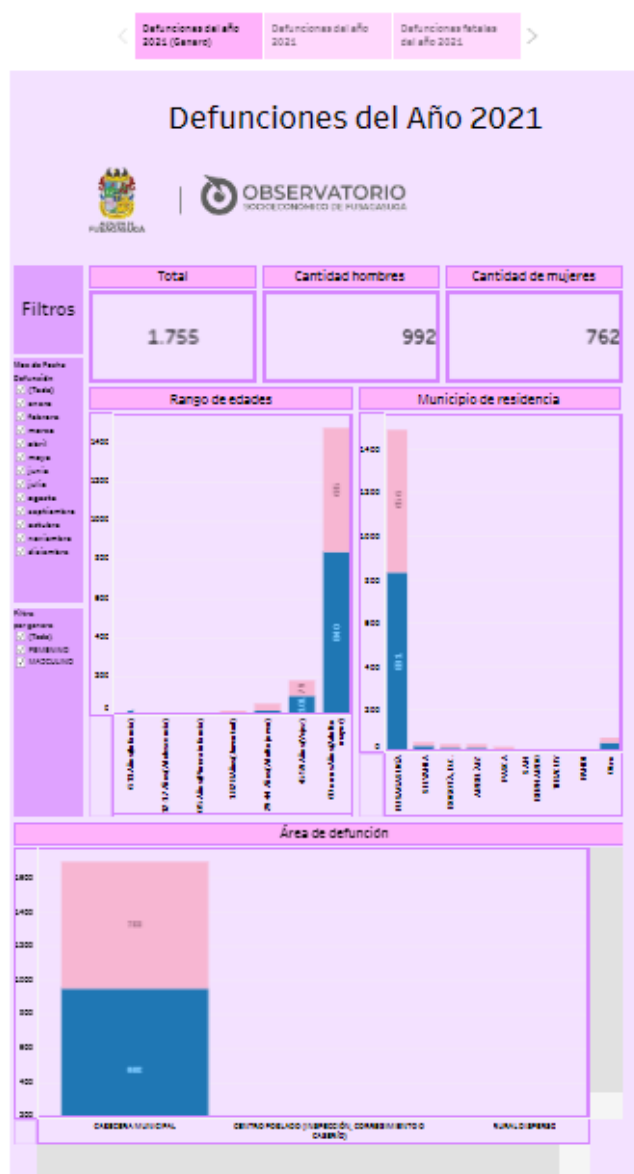
A lo largo del diseño de los diferentes Dashboard se presentaron casos en donde se obtenían más de un producto, por lo que fue necesario crear historias, estas constan de un apartado para agrupar más de un Dashboard, es una manera de llevar un orden para la publicación de los datos, evitando que se tenga que hacer uso de un apartado para cada producto de exposición de información.

Para entender la función de las historias se mostrará un Dashboard que tuvo como resultado tres productos:

- Defunciones del año 2021 (Genero).
- Defunciones 2021
- Defunciones fetales del año 2021

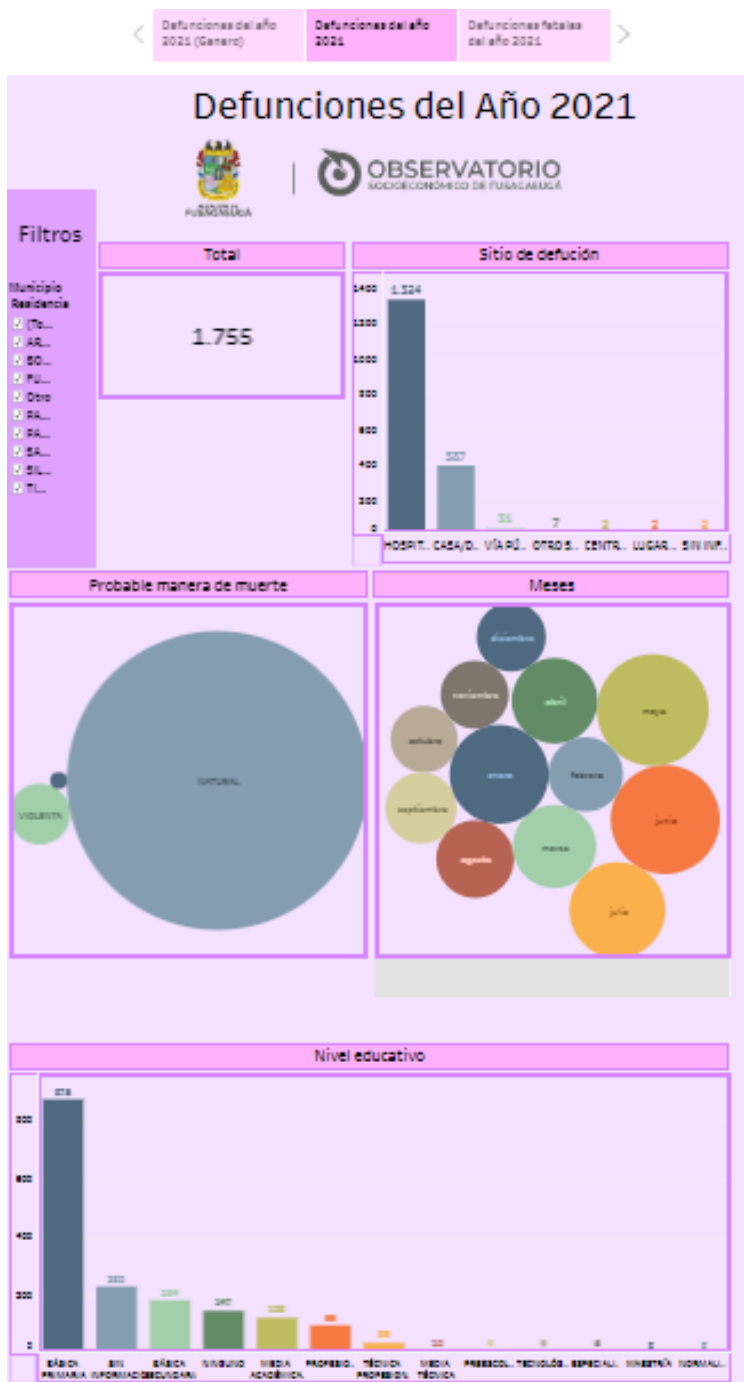
Como se aprecia en las siguientes ilustraciones:

Ilustración 39
Dashboard de defunciones del año 2021 (genero).



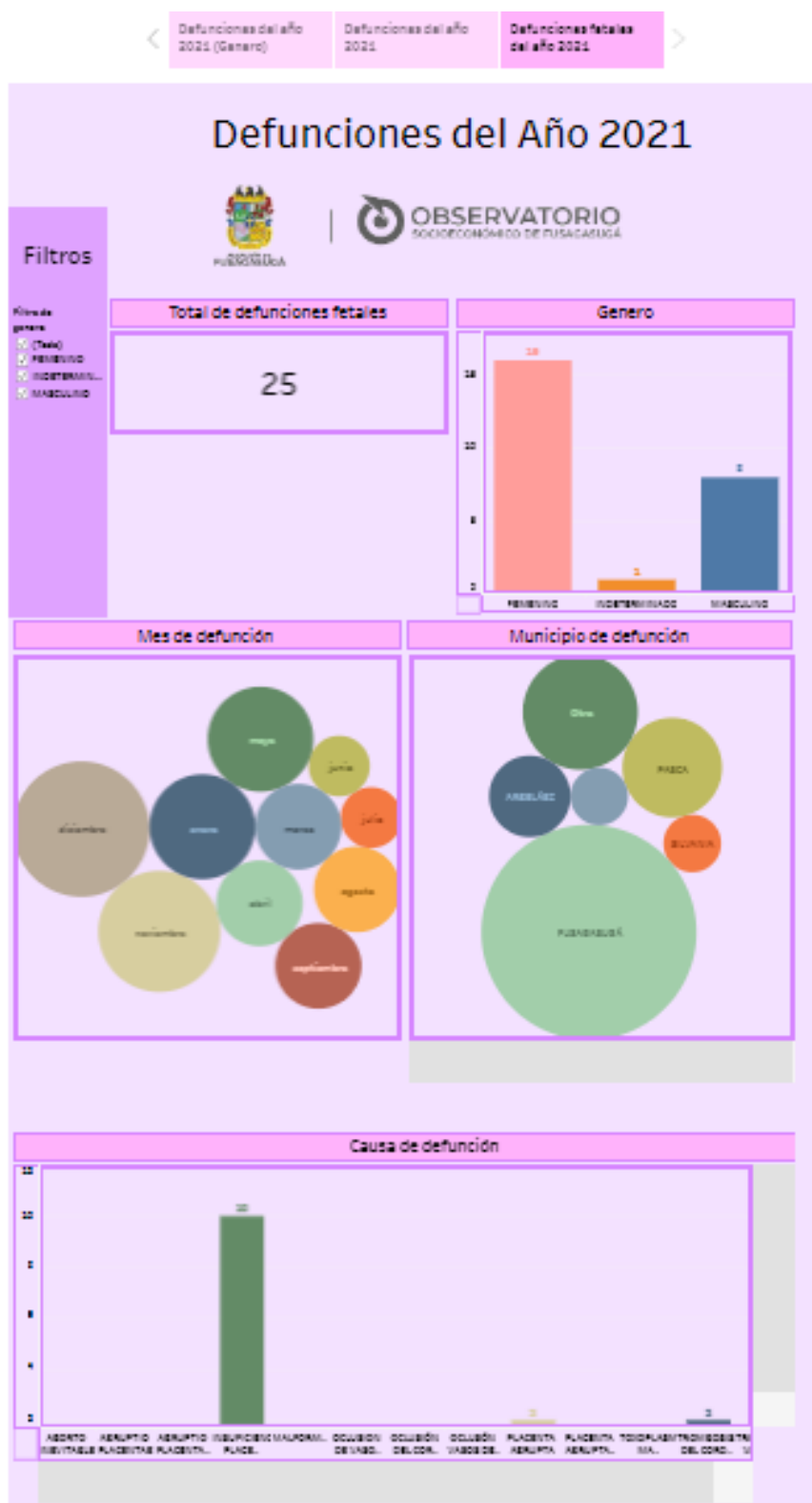
Fuente: Propia.

Ilustración 40
Dashboard de defunciones del año 2021 (Historias).



Fuente: Propia.

Ilustración 41
Dashboard de defunciones del año 2021 (Fetales).



Fuente: Propia.

Observando las ilustraciones anteriores se puede apreciar de mejor manera la función de las historias, evitando el uso innecesario de espacio para la ubicación de varios Dashboard individualmente, lo que permitió una mayor facilidad de la publicación de los datos dentro de la página web del observatorio.

Evaluación y pruebas

Correcciones, evaluación y pruebas

Una vez se ha terminado con la mayor parte del desarrollo de un Dashboard, se obtiene como resultado un primer demo de lo que será el producto final. A lo largo del desarrollo fue necesario realizar una serie de pruebas de cada uno de los apartados del Dashboard, los apartados revisados fueron:

Análisis de concordancia de datos

Fue necesario corroborar que los datos que se están exponiendo dentro de las estadísticas del Dashboard tuvieran concordancia con la información almacenada dentro de los documentos, que facilito en un inicio las dependencias, ya que existía la posibilidad de una inconsistencia de la integridad de los datos, lo que conlleva a la publicación de datos inconsistentes e incorrectos.

se presentó en una serie de tableros con esta inconsistencia antes mencionada, mostrando las estadísticas con datos numéricos diferentes a los de su documento. Para la resolución de estos problemas, basto con hacer una retroalimentación de la información y realizar los cambios de los datos de la estadística, logrando obtener una nueva estadística con los datos reales y así dando cumplimiento a la corrección.

Verificación de filtros y etiquetas

Otra de las pruebas consistía en verificar que la implementación del filtro afectara en todos los campos estadísticos que se tenían previstos, además de esto, se tuvo que verificar que solo afectara a los apartados deseados, ya que a lo largo del proyecto se presentaron en algunas ocasiones, la afectación de las estadísticas en apartados no deseados.

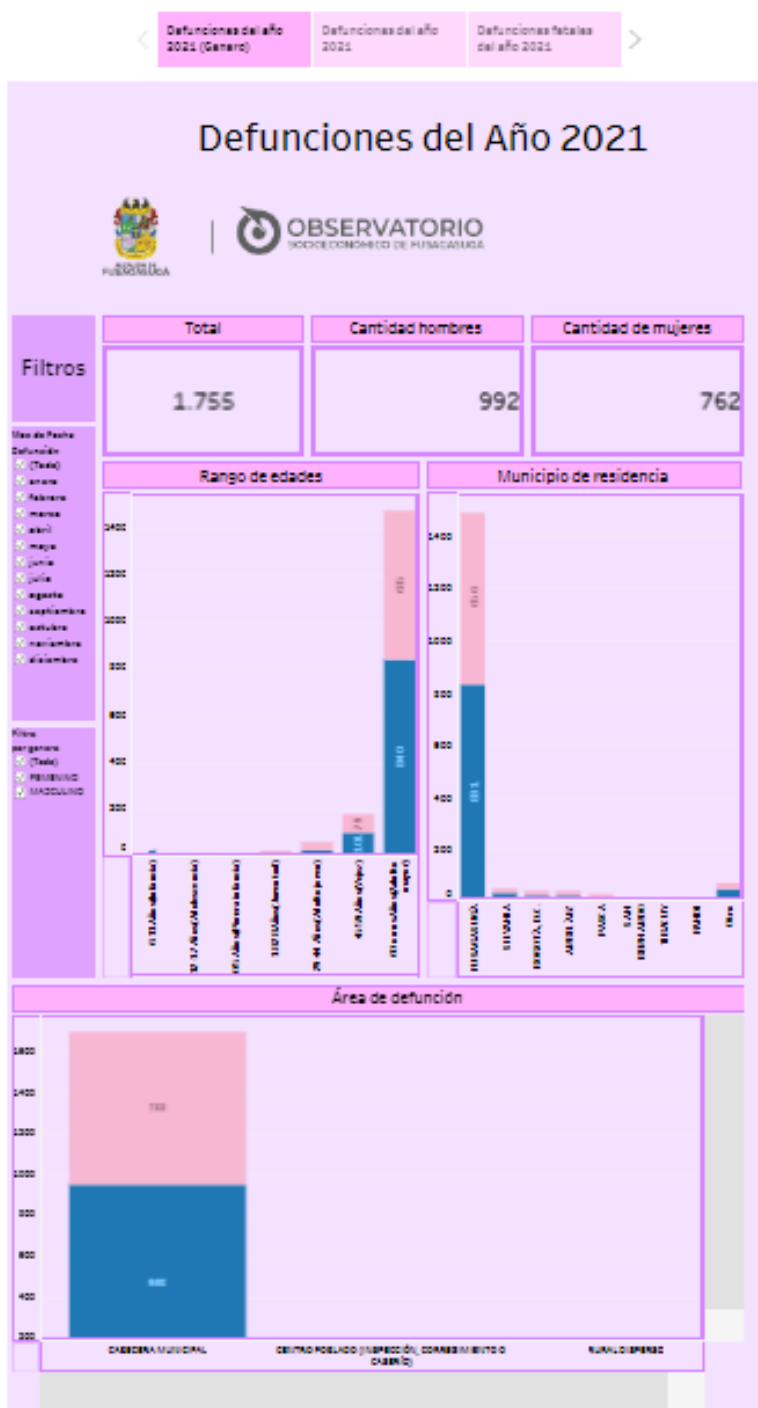
Para corregir este error fue necesario hacer un cambio en la estructuración de la etiqueta para que la filtración de los datos solo se presentara en las estadísticas que se tenían planeadas desde un principio.

Cambios requeridos por el director de proyecto

El ultimo actor en determinar que el desarrollo del producto cumplió con los objetivos propuestos es el jefe de área de Observatorio, se presentaron casos en donde fue necesario hacer el retiro o remplazo de alguno de los componentes del Dashboard, determinando y concluyendo que ese apartado no era necesario incluirlo.

Para ejemplificar un caso donde se requirió del cambio drástico de componentes dentro del Dashboard se expondrá el cambio realizado en el producto resultante del proceso de construcción del documento “defunciones del año 2021”.

Ilustración 42
Dashboard de defunciones del año 2021 demo.

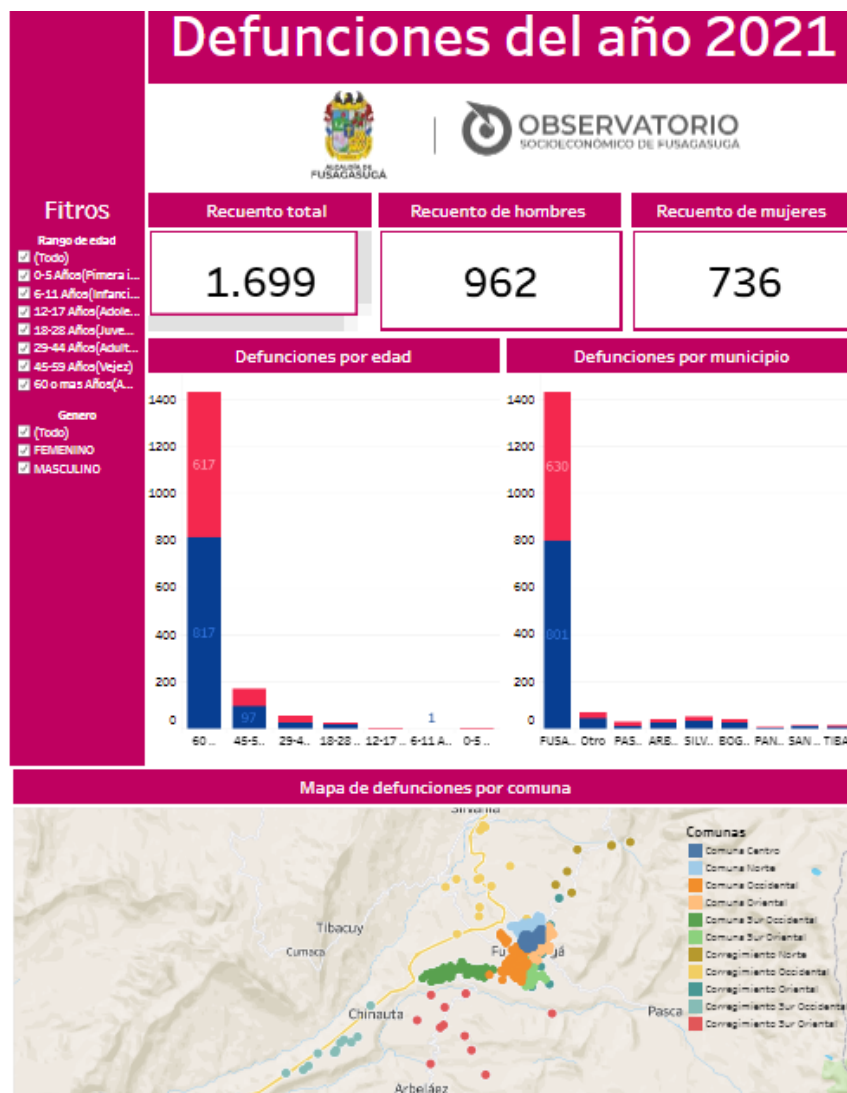


Fuente: Propia.

Y en la ilustración 40 se puede ver su corrección para poder ser aceptada como producto final y ser publicado en la página oficial del observatorio, los cambios realizados constaron de los siguiente:

- Rediseño total del apartado visual del Dashboard, se realizó un cambio en la paleta de colores de cada uno de los ítems que se encuentran dentro del producto, el cambio afecto al apartado de gráficas, filtros, cantidades, títulos e imágenes.
- El cambio más importante dentro de este Dashboard y la razón de su rechazo inicial fue la falta de un mapa de georreferenciación estadística, ya que en un inicio se había implementado una gráfica con información del área de defunción, pero fue reemplazado por un mapa geográfico dividido por comunas, donde se muestra, la cantidad de defunciones por comuna y muestra el punto exacto donde ocurre el fallecimiento.

Ilustración 43
Dashboard de defunciones (Rediseñado).



Fuente: Propia.

Despliegue o puesta en producción

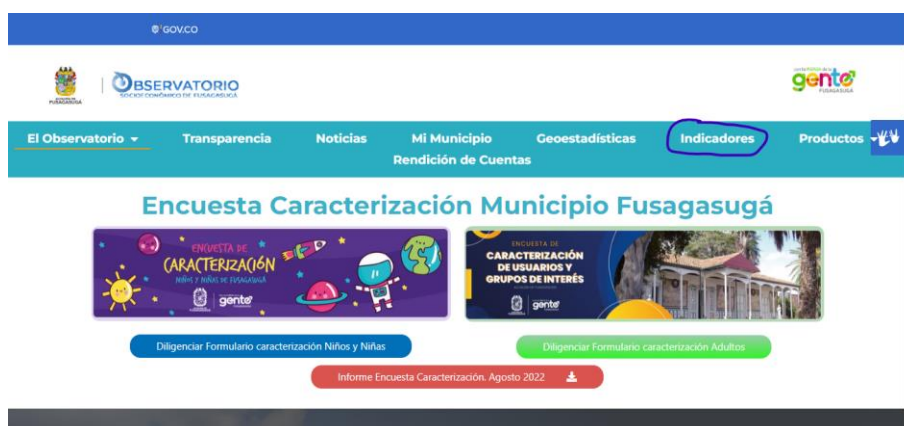
Publicación de los Dashboard en la página web

Todo el proceso anteriormente mencionado culminó en la creación de un tablero de información estadística o en el caso de algunos documentos, culminó en el diseño de una historia, si el producto desarrollado cumplió con las expectativas del equipo de trabajo del Observatorio socioeconómico de la Alcaldía de Fusagasugá, entonces se dio permiso para realizar la publicación del Dashboard en la página oficial del Observatorio socioeconómico de la alcaldía. <https://observatorio.alcaldiafusagasuga.gov.co/>

Dentro de esta página se encuentran un gran número de apartados que fueron diseñados por el equipo de trabajo del Observatorio, se puede encontrar una sección llamada “indicadores”. Donde se encuentran alojados los indicadores de información diseñados para exponer al público en general.

Ilustración 44

Página web Observatorio Fusagasugá.



Fuente: Observatorio Fusagasugá. <https://observatorio.alcaldiafusagasuga.gov.co/>

Dentro del apartado de indicadores se puede encontrar una variedad de opciones de información, donde se exponen por el momento siete categorías de indicadores, cabe resaltar que dentro de cada uno de estos podemos encontrar las subcategorías de los indicadores, así que el total de publicaciones de información varía en cada categoría. Para poder apreciar uno de los Dashboard diseñados y aprobados en el proceso de pasantía, se puede entrar en el apartado de seguridad y convivencia, como se resalta en la ilustración.

Ilustración 45

Indicadores socioeconómicos.



Fuente: Observatorio Fusagasugá. <https://observatorio.alcaldiafusagasuga.gov.co/>

Al momento de entrar al indicador de seguridad y convivencia se puede elegir dos opciones, las cuales se dividen en, violencia de género y hechos delictivos, este último se trata del desarrollo de uno de los Dashboard desarrollados en el proceso de la pasantía, es por ello que se hará el ejemplo con este apartado.

Ilustración 46

Indicadores de seguridad y convivencia



Fuente: Observatorio Fusagasugá. <https://observatorio.alcaldiafusagasuga.gov.co/>

En el segmento de hechos delictivos se encuentran tres opciones de indicadores divididos en seguridad y convivencia 2019, seguridad y convivencia 2020, y seguridad y convivencia 2021. Estos tres indicadores antes mencionados fueron resultado del proceso realizado en el plazo de pasantía, para ejemplificar se mostrará el indicador de los hechos delictivos del año 2021.

Ilustración 47

Indicadores de hechos delictivos.



Fuente: Observatorio Fusagasugá. <https://observatorio.alcaldiafusagasuga.gov.co/>

Dentro del indicador se encuentra el Dashboard de los hechos delictivos ocurridos en la ciudad de Fusagasugá en el año 2021 (<https://observatorio.alcaldiafusagasuga.gov.co/delitos-2021/>) mostrando así el despliegue del producto final de uno de los procesos de construcción de Dashboard.

Ilustración 48

Dashboard hechos delictivos 2021.



Fuente: Observatorio Fusagasugá. <https://observatorio.alcaldiafusagasuga.gov.co/>

Productos desarrollados

En esta sección del documento se presentará un listado de los productos construidos a lo largo de la realización de la pasantía en el observatorio socioeconómico de la alcaldía de Fusagasugá. Algunos de estos productos no han sido publicados dentro de la página oficial del Observatorio a la fecha de realización de este documento, ya que esta característica depende de la rama de desarrollo del Observatorio socioeconómico de la alcaldía de Fusagasugá. Otra cuestión para tener en cuenta dentro del listado es que para cada uno de los indicadores se construyó un documento de análisis de datos (Google Colab) para el desarrollo de los diferentes informes realizados en base a indicadores, llevados a cabo por el encargado del área estadística.

Indicadores de nacimientos en la ciudad de Fusagasugá

Realizado los procedimientos de gestión de los datos del indicador de “Nacimientos en la ciudad de Fusagasugá”, se obtuvo como resultado un documento que contiene el análisis de la información del indicador (Documento Google Colab). Además de esto, se construyeron dos Dashboard que exponen las estadísticas de información del indicador, haciendo énfasis en los niños nacidos vivos de madres menores de edad (14 a 17 años) y el área geográfica en el que se presentaran los nacimientos (rural y urbano).

Hechos delictivos ocurridos en la ciudad de Fusagasugá

En el caso de este indicador, se realizó una gestión de datos a tres indicadores de información que almacenan información similar acerca de hechos delictivos, lo que diferencia a cada uno de estos es el año en el que se presentaron, para este caso en particular, se realizó el trabajo sobre indicadores de los años 2019, 2020 y 2021. Para cada uno de los diferentes indicadores se obtuvieron dos Dashboard, dando un total de seis Dashboard en total, repartidos en tres secciones de la página de la alcaldía de Fusagasugá para cada indicador.

Defunciones en la ciudad de Fusagasugá

Para el indicador de defunciones en la ciudad de Fusagasugá se trabajó con un extenso indicador que contenía 1699 datos almacenados, el producto final de la extracción, transformación y carga de información, resulto en una historia (Espacio de almacenamiento de más de un Dashboard), que contiene tres Dashboard que albergan las estadísticas del indicador, cada uno de los Dashboard hace énfasis en información diferente categorizada en defunciones por género, fecha y manera y por defunciones fetales.

Establecimientos comerciales

Para el caso de los establecimientos comerciales se obtuvo un solo Dashboard, en el que se expone información de los diferentes establecimientos comerciales por medio de categorización de información en las estadísticas, y además de esto, con ayuda de un mapa de calor que expone información geográfica de un establecimiento en particular con ayuda de las etiquetas de información.

Violencia intrafamiliar

Este fue el último indicador (se constaba con dos documentos Excel variables de violencia y tasas de violencia a nivel Fusagasugá, Cundinamarca y Colombia) que se trabajó a lo largo de la realización de la pasantía dentro del Observatorio socioeconómico de la alcaldía de Fusagasugá, el resultado de la gestión de la información fue la construcción de catorce Dashboard, almacenados en dos secciones diferentes.

Para el caso del indicador de tasas de violencia a nivel Fusagasugá, Cundinamarca y Colombia, se construyeron dos Dashboard almacenados dentro de una historia, en el caso del primer Dashboard, se enfatizó en información a nivel Fusagasugá, contando con estadísticas divididas en género y en edad (conteniendo tres categorías niños, niñas y adolescentes, violencia entre otros familiares y violencia hacia el adulto mayor), y en el caso del segundo Dashboard se hizo énfasis en información a nivel Fusagasugá, Cundinamarca y Colombia, conteniendo estadísticas categorizadas de la misma forma que el primer caso.

Para la gestión de información del indicador de variables de violencia intrafamiliar se obtuvieron doce Dashboard, todos estos almacenados en una historia, cada uno de estos Dashboard se dividen por categorías presentes dentro del indicador de información (Documento Excel), las variables que estaban dentro del indicador fueron:

- Violencia por grupo de edad
- Violencia por ciclo vital
- Violencia por escolaridad
- Violencia por estado civil
- Presunto agresor
- Actividad realizada durante el hecho
- Factor desencadenante
- Topografía de las lesiones
- Escenario del hecho
- Zona del hecho
- Mes del hecho
- Día del hecho

Recomendaciones

Dentro del proceso de construcción de las estadísticas, fue necesario realizar una transformación completa de todos los documentos que almacenaban la información de los indicadores. Se recomienda implementar el uso de bases de datos (por ejemplo Oracle, MySQL o Microsoft SQL Server), como medio de almacenamientos de información; y a su vez, desempeñar la creación de un aplicativo software para todo el manejo y estructuración de la información dentro de los procesos de obtención de información. De esta forma, se trabajará con un medio general para cada una de las oficinas de la alcaldía, y se evitara problemáticas con los diferentes formatos de información que manejan las oficinas internas con las que trabaja el observatorio.

Se recomienda implementar la utilización de metodologías de manejo de información para los datos proporcionados por las diferentes dependencias, ya que, a lo largo de la realización del proyecto fue necesario realizar una variedad de correcciones en los documentos que almacenaban la información de los indicadores, estas correcciones mencionadas, son propiedades a tener en cuenta en las dimensiones de la calidad de datos, de ahí que, si se implementa el uso de este medio de verificación de calidad, se podrá ejercer una mayor cantidad de procesos en un menor tiempo y con una mayor rectitud, evitando la necesidad de corrección por no concordancias en los datos.

Por ultimo se recomienda hacer uso de una herramienta para el manejo de dato ETL (se recomienda apache airflow), este tipo de herramientas se encargará de la adecuación de los diferentes datos, a implementar el uso de la herramienta dentro de la fuente de información, se obtendrá una serie de datos organizados para trabajar de una manera homogénea, corrigiendo defectos en las cantidades de variables y los errores dentro de los datos, lo que concluye en un documento con calidad en sus datos.

Conclusiones

Los datos obtenidos por las diferentes dependencias con las que tiene convenio la Alcaldía de Fusagasugá fueron sometidos a un exhaustivo estudio, seleccionando los datos de mayor impacto informativo, exceptuando cada dato de identificación personal, transformando los datos en pro de una mejor organización de información. Terminado el proceso se pudo iniciar con el desarrollo de las estadísticas y posteriormente la creación del tablero, es aquí donde se cumplió el objetivo de la adecuación de la información, ya que se logró presentar los datos estadísticos de una manera intuitiva y de fácil comprensión, representándose por medio de una interfaz gráfica expuesta en la web.

Dentro del diseño del proyecto se implementó una tecnología de georreferenciación de datos informativos, por medio del diseño de mapas estadísticos, tomando como base datos de referenciación topológica. Gracias a la herramienta de gestión de mapas de Tableau, se logró diseñar una serie de mapas interactivos que lograban exponer de una manera diferente e innovadora, información específica, esta socialización de información se logró en conjunto con las etiquetas de información, con las que contaban cada uno de los mapas, haciendo que un usuario ya sea ciudadano común o trabajador de alguna dependencia de la alcaldía, solo tenga que colocar el cursor de su dispositivo para poder acceder a estos datos de interés social y económico en el municipio.

La información expuesta por las estadísticas se categorizo de manera sectorial, en petición de la Alcaldía de Fusagasugá, esto hace que los análisis sectoriales y la comprensión de información sea más eficaz y sencilla de realizar para las oficinas que componen la entidad, además de esto, se realizó una intervención en la búsqueda de información, convirtiéndola en un objeto más sencillo de ubicar por parte de un usuario, minimizando la complejidad de búsqueda de información sectorial y facilitando su entendimiento, esto en conjunto realizado con la página web, la cual brinda el espacio de alojamiento de la información estadística, (creada por la instancia de desarrollo de software del observatorio socioeconómico).

Referencias bibliográficas

Mintic. Guía para el uso y aprovechamiento de Datos Abiertos en Colombia.

<https://herramientas.datos.gov.co/sites/default/files/Guia%20de%20Datos%20Abiertos%20de%20Colombia.pdf>

Baque (2014). análisis de la base de datos del hospital universitario de guayaquil para la creación de reportes gráficos con la herramienta tableau.

<http://repositorio.ug.edu.ec/bitstream/redug/10229/1/PTG-628%20Baque%20Carrasco%20Kevin%20Antonio.pdf>

Villavicencio (2015). Análisis de los pacientes ingresados en ucín vs el número de nacimientos vivos del hospital universitario de guayaquil, para la generación de reportes utilizando la herramienta tableau public y sql server 2008.

<http://repositorio.ug.edu.ec/bitstream/redug/10113/1/PTG-829%20Villavicencio%20Ponce%20Evelyn%20Patricia.pdf>

Silva (2022). Desarrollar un conjunto de tableros visuales de Business Intelligence (BI) mediante la herramienta de Tableau que permita fortalecer la interpretación de la información de deserción estudiantil en la Universidad Técnica del Norte.

<http://repositorio.utn.edu.ec/bitstream/123456789/12217/2/04%20ISC%20609%20TRABAJO%20GRADO.pdf>

Cabanillas (N.d). Calidad de datos.

https://openaccess.uoc.edu/bitstream/10609/140887/2/Gobierno%20del%20dato_M%C3%B3dulo%203_Calidad%20de%20datos.pdf

Espinosa. (2020). Aplicación de metodología CRISP-DM para segmentación geográfica de una base de datos pública. <https://www.scielo.org.mx/pdf/iit/v21n1/1405-7743-iit-21-01-e00008.pdf>

García (2018). Aplicación de la metodología CRISP-DM a la recolección y análisis de datos georreferenciados desde Twitter. <https://repository.unimilitar.edu.co/bitstream/handle/10654/20099/GarciaVelezGustavoAdolfo2018.pdf?sequence=3&isAllowed=y>

Gómez (2018). Conector web para la integración de datos abiertos colombianos con la herramienta Tableau. http://repositorio.itm.edu.co/bitstream/handle/20.500.12622/1719/Rep_Itm_pre_Gomez.pdf?sequence=1&isAllowed=y

Escobar (Universidad Javeriana - 2016). Estandarización de los reportes del análisis de resultados de las pautas publicitarias digitales de la agencia MediaCom Colombia para sus clientes a través del software Tableau. <https://repository.javeriana.edu.co/bitstream/handle/10554/19672/EscobarHuertasMariaAlejandra2016.pdf?sequence=1&isAllowed=y>

Camacho & Dávila (2012). Georreferenciación de documentos cartográficos para la gestión de Archivos y Cartotecas. <https://www.ign.es/web/resources/docs/IGNCnig/CTC-Ibercarto-V-Georreferenciacion.pdf>

Velázquez (N.d) Conceptos básicos de estadística. <https://centrogeo.repositorioinstitucional.mx/jspui/bitstream/1012/157/1/13-Conceptos%20B%C3%A1sicos%20de%20Estad%C3%ADstica%20-%20Diplomado%20en%20An%C3%A1lisis%20de%20Informaci%C3%B3n%20Geoespacial.pdf>

Galán (2015). Aplicación de la metodología crisp-dm a un proyecto de minería de datos en el entorno universitario. [https://e-](https://e-archivo.uc3m.es/bitstream/handle/10016/22198/PFC_Victor_Galan_Cortina.pdf)

[archivo.uc3m.es/bitstream/handle/10016/22198/PFC_Victor_Galan_Cortina.pdf](https://e-archivo.uc3m.es/bitstream/handle/10016/22198/PFC_Victor_Galan_Cortina.pdf)

Portilla (2021). Diseño, creación e implementación de un Dashboard para el seguimiento de peticiones de servicios en el área de mantenimiento, planeación y control de FORD

España. [https://riunet.upv.es/bitstream/handle/10251/174864/Portilla%20-](https://riunet.upv.es/bitstream/handle/10251/174864/Portilla%20-%20Diseno%20creacion%20e%20implementacion%20de%20un%20Dashboard%20para%20el%20seguimiento%20de%20peticiones%20de%20...pdf?sequence=1)

[%20Diseno%20creacion%20e%20implementacion%20de%20un%20Dashboard%20para%20el%20seguimiento%20de%20peticiones%20de%20...pdf?sequence=1](https://riunet.upv.es/bitstream/handle/10251/174864/Portilla%20-%20Diseno%20creacion%20e%20implementacion%20de%20un%20Dashboard%20para%20el%20seguimiento%20de%20peticiones%20de%20...pdf?sequence=1)

SAS (2022). SAS Inteligencia de Cliente 360: Guía del Usuario.

<https://documentation.sas.com/api/docsets/cintug/22.11/content/cintug.pdf?locale=es#na-meddest=personally-identifiable-info>

Magal & Rincón (2019). Plan de gestión de calidad de datos para mejorar la oportunidad y pertinencia de la información de la oferta institucional en la dirección de apropiación del

ministerio tic. <https://bdigital.uexternado.edu.co/server/api/core/bitstreams/6179eea0-6381-4ebb-bf0b-473a09298914/content>

Tesel. Convertir direcciones a coordenadas de latitud y longitud con Google Sheets y Google

Maps (Geocoding). <https://tesel.mx/convertir-direcciones-a-coordenadas-de-latitud-y-longitud-con-google-sheets-y-google-maps-geocoding-7928/>