Análisis de Datos sobre las causas de la Deforestación en el Departamento del Chocó (2014-2021)

Luis Carlos Abril Betancourt, Yudy Viviana Bran Sánchez, Ingrith Lorena Macias Suarez y Jeison Alexander Hilarion Parra

Especialización en Big Data, Fundación Universitaria Compensar

Proyecto fin de Especialización en Big Data

Omar Alexander Leon García

29 de mayo de 2024

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

Análisis de datos sobre las Causas y Extensión de la Deforestación en el Departamento del Chocó (2014-2021)

Luis Carlos Abril Betancourt
Yudy Viviana Bran Sánchez
Ingrith Lorena Macias Suarez
Jeison Alexander Hilarion Parra

Trabajo de grado presentado como requisito parcial para optar al título de:

Especialista en Big Data.

Director (a):

Ph.D. Omar Alexander León García

Grupo de Investigación:

Grupo de investigación GIIS

Fundación Universitaria Unipanamericana – Compensar

Facultad de Ingeniería, Especialización en Big Data

Bogotá D.C, Colombia

2024

Resumen

La deforestación en el Chocó representa una amenaza crítica para la biodiversidad y el equilibrio climático de la región. Para comprender mejor la complejidad de este problema, se requiere un análisis exhaustivo de datos. Utilizando la metodología SEMMA, se examinaron datos recopilados entre 2014 y 2021, revelando que la deforestación es principalmente resultado de actividades humanas como el crecimiento urbano, la agricultura y la minería, lo que ha provocado una pérdida significativa de biodiversidad y cambios climáticos por lo cual para visualizar y comprender mejor estos hallazgos, se empleó la herramienta Power BI, creando tableros interactivos que muestran la distribución geográfica y el análisis temporal de la deforestación. Estas visualizaciones ayudan a identificar tendencias y patrones significativos, proporcionando una base sólida para la toma de decisiones informadas y la implementación de medidas efectivas para abordar este problema ambiental en el Chocó.

Palabras clave:
Deforestación, Metodología SEMMA, Power BI, Análisis de datos, Departamento del Chocó

Abstract

La deforestation in Chocó represents a critical threat to the biodiversity and climatic balance of the region. To better understand the complexity of this issue, a comprehensive data analysis is required. Using the SEMMA methodology, data collected between 2014 and 2021 were examined, revealing that deforestation is mainly the result of human activities such as urban growth, agriculture, and mining, leading to significant loss of biodiversity and climate changes. To visualize and better understand these findings, the Power BI tool was employed, creating interactive dashboards that display the geographic distribution and temporal analysis of deforestation. These visualizations help identify significant trends and patterns, providing a solid foundation for informed decision-making and the implementation of effective measures to address this environmental issue in Chocó.

Keywords:
Deforestation, Chocó, Data analysis, SEMMA methodology, Power BI

Contenido

RESUMEN ..... **III**

PALABRAS CLAVE: III

ABSTRACT ..... **V**

KEYWORDS: V

INTRODUCCIÓN **1**

ANTECEDENTES 2
JUSTIFICACIÓN 4
FORMULACIÓN DEL PROBLEMA 5
OBJETIVO GENERAL 5
OBJETIVOS ESPECÍFICOS 5
ALCANCES 5
LIMITACIONES 6

1. MARCO TEÓRICO **7**

1.1 DEFORESTACIÓN: CAUSAS Y CONSECUENCIAS 7
1.1.1 Causas de Deforestación *7*
1.1.2 Consecuencias de Deforestación *9*
1.2 ANÁLISIS DE DATOS Y VISUALIZACIONES CON POWER BI 10
1.2.1 Recopilar y preparar datos *10*
1.2.2 Desarrollar visualizaciones significativas *11*
1.2.3 Interpretación de resultados *11*
1.2.4 Ciclo de iteración y mejora continua *12*

2. DISEÑO METODOLÓGICO **13**

2.1 FASE I - SAMPLE (MUESTREO) 14
2.2 FASE II - EXPLORE (EXPLORACIÓN) 14
2.3 FASE III - MODIFY (MANIPULACIÓN) 14
2.4 FASE IV - MODEL (MODELA) 14
2.5 FASE V – ASSESS (VALORACIÓN) 15

3. RESULTADOS **17**

4. DISCUSIÓN **23**

5. CONCLUSIONES Y RECOMENDACIONES **24**

5.1 CONCLUSIONES 24
5.2 RECOMENDACIONES 25
BIBLIOGRAFÍA 27

Introducción

La deforestación en el Departamento del Chocó entre 2014 y 2021 es muy importante y preocupante por su impacto en el medio ambiente, la biodiversidad y las comunidades locales. Se realizará un análisis de datos enfocado en las causas y la magnitud de este fenómeno en la región. El Chocó, ubicado en la costa pacífica de Colombia, es reconocido por su riqueza biológica y su diversidad étnica y cultural. Sin embargo, enfrenta desafíos significativos en términos de deforestación, que afecta los ecosistemas naturales, la calidad del agua y el hábitat de numerosas especies.

Durante el período mencionado, se ha observado un aumento alarmante en la tasa de deforestación en el Departamento del Chocó, lo que ha generado preocupación a nivel local, nacional e internacional. Según con el EDEAM solo para el 2021 el Choco presento una perdida aproximada de 295.000 hectáreas, dando así la relación con el Global Forest Watch esto corresponde a 159 Millones de toneladas de CO2. Y para el año 2019 problemáticas causadas por las acciones del ser humano fue aproximadamente sobre 177.644 hectáreas (Forero,2023).

Este análisis pretende comprender las causas de este fenómeno, desde la expansión de la agricultura y la ganadería hasta la minería ilegal y la deforestación a gran escala. Además de determinar las causas, se considerará el alcance de la deforestación tanto en las áreas afectadas como en las relacionadas e impactos en la biodiversidad, así como en las comunidades indígenas y afroamericanas cuyos medios de vida y culturas dependen de los recursos naturales.

El proyecto se enfoca en abordar la falta de análisis de datos exhaustivos sobre las causas de la deforestación en el Departamento del Chocó, Colombia. A pesar de su impacto significativo en el medio ambiente, la biodiversidad y las comunidades locales, existe una carencia de

información detallada sobre las causas específicas y sus interrelaciones. Esta falta de comprensión dificulta la implementación de medidas efectivas para mitigar la deforestación y proteger los ecosistemas naturales. Por lo tanto, el proyecto busca llenar este vacío mediante un análisis riguroso de datos que identifique las principales causas de la deforestación, sus conexiones y sus impactos en la región. Esto proporcionará una base sólida para el diseño e implementación de estrategias de conservación y desarrollo sostenible que aborden de manera integral este problema y promuevan un uso responsable de los recursos naturales en el Chocó.

Con este análisis pretendemos identificar y analizar las causas y tendencias de la deforestación en el Departamento del Chocó para concienciar a la importancia de la conservación de los bosques. Se propondrá el desarrollo de visualizaciones utilizando la herramienta Power BI, que permitan explorar y comprender intuitivamente los datos sobre la deforestación en el Chocó. Estas visualizaciones no solo facilitarán la identificación de patrones y tendencias, sino que también servirán como herramientas poderosas para comunicar los hallazgos a una variedad de audiencias, desde científicos y tomadores de decisiones hasta comunidades locales y partes interesadas.

Antecedentes

El Chocó es una de las regiones más biodiversas del mundo que cuenta con ecosistemas únicos como los bosques tropicales húmedos y los manglares. La conservación de estos ecosistemas es crucial no solo a nivel local, sino también globalmente, debido a su contribución a la mitigación del cambio climático y la conservación de la biodiversidad, este departamento cuenta con 327.000 Ha equivalente al 85% de la totalidad de la cobertura de manglares en el país y está siendo amenazada, principalmente por el cambio de uso del suelo para agricultura, ganadería, cultivos ilícitos, proyectos de infraestructura y minería (Imeditores,2009, cap.3) .

En el análisis realizado por Restrepo López acerca de los elementos vinculados con la disminución de la cobertura forestal en los municipios de Riosucio y Río Quito, Chocó, Colombia, se determinó que, a nivel nacional, los principales desencadenantes de la deforestación son tres: la densidad poblacional, la forma en que se utiliza el suelo y la inclinación del terreno (Lopez,2022). Estos autores exponen que las zonas rurales de baja elevación, con poca ganadería, y de bajas pendientes se asocian con una alta deforestación, lo

que sugiere que la pérdida de bosques ocurre en mayor extensión en áreas remotas con frentes activos de colonización y topografía plana. Sin embargo, el comportamiento de estas variables varía entre regiones. Este estudio muestra que variables como la pendiente, la distancia a áreas protegidas, áreas deforestadas, vías primarias, títulos y solicitudes mineras, centros 30 urbanos, vías secundarias, cuerpos de agua y ríos son importantes para explicar la deforestación local.

David Forero, en su trabajo de investigación deforestación en el departamento del chocó causas, datos y enfoques asociados al concepto de seguridad multidimensional concluyo La deforestación en el Chocó como fenómeno, es una problemática de impactos ambientales y territoriales. Para 2021, el aumento de hectáreas deforestadas en el departamento fue de 12,6%. Las principales causas son los cultivos de hoja ilegal de coca, la explotación ilícita de yacimientos mineros y los incendios forestales. Esta investigación se centró en la minería y los cultivos. El ejercicio de análisis de datos condujo a la delimitación de tres municipios con mayor impacto y/o afectación: Unguía, Riosucio y el Cantón San Pablo. Los tres territorios poseen características micro etnográficas complejas, específicamente en materia económica (Forero,2023).

La deforestación en el departamento del chocó: causas, datos y enfoques asociados al concepto de seguridad multidimensional, se concluyó que la deforestación en el Chocó como fenómeno, es una problemática de impactos ambientales y territoriales. Para 2021, el aumento de hectáreas deforestadas en el departamento fue de 12,6% (Telles,2021) teniendo como las principales causas los cultivos de hoja ilegal de coca, la explotación ilícita de yacimientos mineros y los incendios forestales. Esta investigación se centró en la minería y los cultivos, por lo cual el ejercicio de análisis de datos condujo a la delimitación de tres municipios con mayor impacto y/o afectación: Unguía, Riosucio y el Cantón San Pablo; los tres territorios poseen características micro etnográficas complejas, específicamente en materia económica por consiguiente los niveles de pobreza multidimensional registrados, aunados a unos bajos índices de desarrollo humano y un nivel promedio de necesidades básicas insatisfechas de 55,6%, son, desde la perspectiva del autor, acelerantes en la problemática de deforestación

La anterior es una hipótesis que se planteó una vez finalizó el análisis correlacional de datos presentados en los impactos ambientales y territoriales. Entre los impactos ambientales más relevantes no solo estuvo el número de hectáreas deforestadas, también se halló la pérdida

constante de hectáreas de bosque húmedo (1% hasta 2021) y reservas forestales con ecosistemas especiales, en cuanto a los impactos territoriales, la situación es un poco más compleja. Los datos registrados en el Geovisor de indicadores económicos del DANE (2022) advirtieron que, primero, los municipios con mayor deforestación también son territorios con un promedio IPM de 64,2%. Significa que las poblaciones habitantes son altamente vulnerables por un amplio margen de pobreza extrema y pobreza monetaria y los territorios con más deforestación están ligados a una restricción de acceso de servicios básicos como educación, saneamiento o cobertura médica no menor al promedio de 48,3%.

Frente a este tipo de escenarios en los que el indicador pobreza es común y transeccional, surgen fenomenologías delictivas complejas quizá poco estudiadas con la óptica del concepto de seguridad multidimensional, precisamente con el enfoque de seguridad humana, pero esta es una conclusión que se expone al final. Primero hay que explicar las deducciones asociadas a las causas del fenómeno. (Forero,2023).

Justificación

La deforestación en la provincia del Chocó es un problema ambiental y socioeconómico muy urgente que requiere atención urgente para implementar medidas de mitigación específicas. La construcción de este estudio se justifica por el impacto ambiental, que conlleva consecuencias devastadoras para los ecosistemas naturales, la biodiversidad y los servicios ecosistémicos de la provincia del Chocó, y por impactos socioeconómicos para la sociedad, y afecta a las comunidades locales dependientes de los bosques. recursos para su supervivencia.

Además, se necesita un análisis de datos para determinar las causas de la deforestación en el Chocó, proporcionando una imagen precisa y detallada de las actividades humanas y de vida silvestre que contribuyen a la deforestación, comprendiendo las causas de la deforestación, es posible desarrollar estrategias y políticas más efectivas para proteger los bosques y utilizar sosteniblemente los recursos naturales en el Chocó, facilitando el monitoreo continuo de las tendencias de la deforestación en el tiempo permitiendo así evaluar la efectividad de las acciones tomadas y, si es necesario, ajustar estrategias para lograr resultados ambientales positivos.

Formulación del Problema

La deforestación en el Departamento del Chocó es un problema persistente que requiere una comprensión más profunda de sus causas principales y el alcance del daño ambiental asociado. Aunque se ha observado una pérdida significativa de cobertura forestal en la región durante los años 2014 a 2021, aún se carece de un análisis detallado que identifique las causas subyacentes y cuantifique el alcance exacto de la deforestación. Por lo tanto, ¿Cómo el análisis de datos facilita la identificación y toma de decisiones sobre las causas de la deforestación en el departamento del Chocó?

Objetivo General

Construir un tablero de control que permita visualizar los datos sobre las causas de la deforestación en el Chocó durante el período comprendido entre los años 2014 y 2021 mediante la herramienta Power BI.

Objetivos Específicos

- Obtener los datos desde el catálogo de datos abiertos de Colombia que permitan identificar las causas sobre la deforestación en el Chocó durante el período especificado
- Identificar la metodología SEMMA para el desarrollo del tablero en Power BI que permite relacionar las variables de deforestación del choco y visualizar de manera clara y comprensible la evolución temporal de las causas.
- Validar las funcionalidades interactivas en el tablero de control de Power BI que permitan a los usuarios identificar las causas de deforestación del Choco de manera dinámica

Alcances

- Se realiza un estudio centrado en identificar y analizar las causas principales de la deforestación en el Departamento del Chocó durante el período de estudio.
- Se tomarán los datos reportados por el ministerio de ambiente para determinar la cantidad de Hectáreas afectadas por la deforestación en cada año

- Se realizará un estudio para lograr identificar si existe alguna tendencia de deforestación en el transcurso del tiempo en los municipios del departamento de Choco.
- Se llevará un análisis de los municipios para detectar cual es la razón principal de cada uno y así proponer medidas de mitigación.

Limitaciones

- Se intenta conseguir la mayor cantidad de información confiable de las fuentes correspondientes, teniendo en cuenta que puede existir limitación entre el espacio reportado
- En base a la información pueden existir datos donde no fue posible la detección de la causa de la deforestación, aun así, esta se contemplará como desconocida para tener el total de las hectáreas afectadas.

1. Marco teórico

Se analizará el problema de la deforestación definiendo sus causas y/o consecuencias y los diferentes frentes que puede llegarse a ver involucrados. También se realizará la definición de la herramienta BI la cual es una herramienta muy utilizada para el procesamiento de datos, que permitirá la visualización de los datos analizados y la metodología de investigación Semma utilizada para el desarrollo de la investigación desarrollando la importancia del análisis de datos y las visualizaciones para abordarlo de manera efectiva.

1.1 Deforestación: Causas y Consecuencias

Para dar paso a las causas y consecuencias de la deforestación, se analizará el principio de este y es definir que es la deforestación. La deforestación una causa que conlleva a un problema climático que es causado en varias ocasiones por factores naturales, pero muchas más por procedimientos realizados por los seres humanos. En este apartado se examinarán las causas de la deforestación y las consecuencias ambientales, sociales y económicas del fenómeno (Natalia & Camilo, 2022).

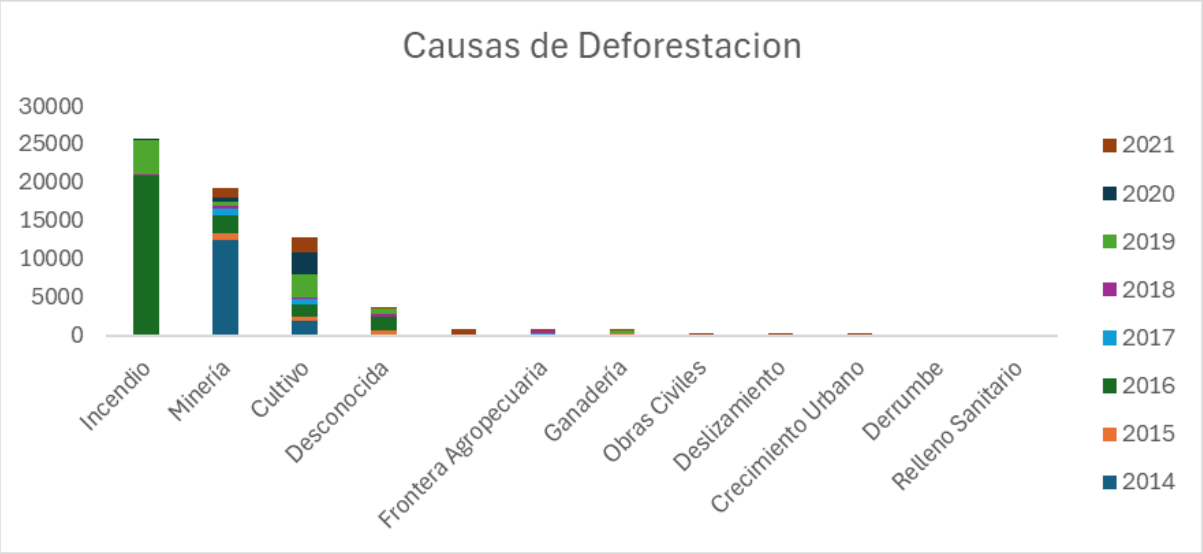
1.1.1 Causas de Deforestación

El crecimiento urbano, caracterizado por la expansión de áreas urbanas y la construcción de infraestructuras como carreteras, edificios y servicios públicos, ha sido un factor clave en la transformación del paisaje natural. La conversión de áreas forestales en terrenos agrícolas para la producción de cultivos comerciales, también ha contribuido significativamente a la deforestación. Además, fenómenos naturales como los derrumbes y deslizamientos, que pueden ser exacerbados por la deforestación, resultan en la pérdida de suelo y vegetación, incrementando la vulnerabilidad de los ecosistemas. La expansión de la frontera agropecuaria y la ganadería extensiva han llevado a la degradación y pérdida de bosques, mientras que

actividades humanas como la minería y la construcción de obras civiles alteran considerablemente el paisaje y la cobertura forestal. Los incendios forestales, a menudo provocados por la quema de vegetación para despejar terrenos, y el uso de áreas forestales para rellenos sanitarios, también degradan el ecosistema. Cada uno de estos factores contribuye de manera integral a la deforestación y la pérdida de biodiversidad, subrayando la necesidad de estrategias de manejo sostenible del territorio Asprilla (2019).

En la Figura 1 evidenciamos la cantidad de Hectáreas que han sido consumidas en el pasar de los años por cada una de las causas presentadas en el sistema de datos abiertos, teniendo como referencia que la causa con más Hectáreas son Incendios en el año 2016 y minería en el año 2020, aun así, esta segunda se ha presentado en más años con una cantidad constante de Hectáreas consumidas (Natalia & Camilo, 2022).

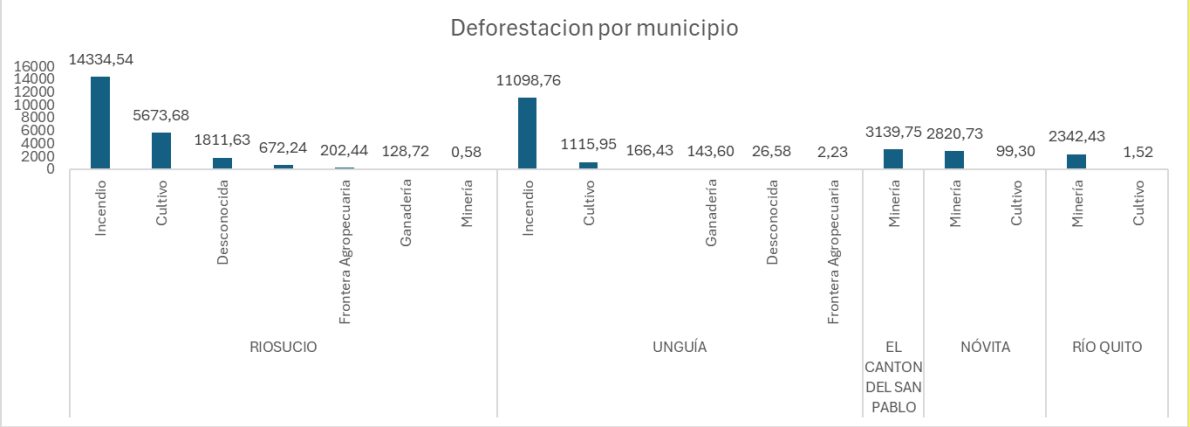
Figura 1 Causas de deforestación



Fuente: Información recuperada de Sistema de Datos Abiertos de Colombia y Corporación Autónoma Regional para el Desarrollo Sostenible del Chocó (2022).

En la Figura 2 observamos como los municipios de Riosucio y Ungia han sufrido de deforestación principalmente por los incendios presentados, y los cultivos ilegales que se han tenido que erradicar afectando la zona, en cuanto El canton del San Pablo, Novita y Rioquito, presentan un enfoque principal en Minería que cada vez va afectando más a los municipios y a las zonas boscosas (Natalia & Camilo, 2022).

Figura 2 Top 5 de Deforestación por municipio



Fuente: Información recuperada de Sistema de Datos Abiertos de Colombia y Corporación Autónoma Regional para el Desarrollo Sostenible del Chocó (2022).

1.1.2 Consecuencias de Deforestación

La deforestación tiene múltiples impactos negativos, entre los cuales destaca la pérdida de biodiversidad, ya que la destrucción de hábitats naturales conlleva la desaparición de numerosas especies vegetales y animales. Este proceso también contribuye significativamente al cambio climático debido a la emisión de gases de efecto invernadero derivados de la eliminación de la cobertura forestal. La erosión del suelo y la degradación ambiental son consecuencias directas de esta pérdida de bosques, aumentando el riesgo de que el suelo se deteriore y pierda su fertilidad. Además, la deforestación tiene un impacto severo en las comunidades indígenas y locales, que dependen de los bosques para su subsistencia, viéndose afectadas tanto cultural como económicamente. Asimismo, la reducción de la cobertura forestal disminuye la capacidad de los ecosistemas para proporcionar servicios esenciales, como la regulación del clima, el control de inundaciones y la provisión de agua limpia, lo que agrava los problemas ambientales y socioeconómicos (Asprilla, 2019).

En la figura 3 evidenciamos que desde el 2018 se ha presentado un aumento en la cantidad de Hectáreas deforestadas, teniendo presente que se quitó la causal de incendio por lo cual, se centraría en minería y cultivos ilícitos que han ido afectando al departamento, siendo este un factor en el cual toca prestar atención principal para erradicar dichas acciones en contra del ambiente (Asprilla,2019).

Figura 3 Deforestación por año sin la causal de incendio



Fuente: Información recuperada de Sistema de Datos Abiertos de Colombia y Corporación Autónoma Regional para el Desarrollo Sostenible del Chocó (2022).

1.2 Análisis de Datos y Visualizaciones con Power BI

Power BI ".es una colección de servicios de software." (Davidiseminger, 2024). Es una herramienta de Ciencia de datos que brinda acceso a datos fácil y seguro, ya que le permite descargar o conectarse a varias fuentes de datos; como Excel, Internet, API, base de datos, etc.

Permite limpiar, transformar, analizar y luego crear esta información en forma gráfica mediante el desarrollo de informes y paneles, lo que le permite visualizar los datos procesados, lo que hace que el material esté más organizado y sea más comprensible para los consumidores (Diaz, 2022).

Los siguientes pasos detallan el proceso a seguir de recopilación, limpieza y visualización de datos, con el objetivo de utilizarlos en la herramienta Power BI con el fin de extraer la informacion del conjunto de datos y representarlos de manera mas comprensible ante el usuario.

1.2.1 Recopilar y preparar datos

Recopilación de Datos: Este paso implica reunir información relevante sobre dispositivos y otros factores relacionados con la deforestación en el departamento de Chocó. Esto puede

incluir datos del Ministerio de Medio Ambiente, datos satelitales, informes de organizaciones no gubernamentales, entre otros (Diaz, 2022).

Limpieza de Datos: Durante este proceso, se identifican y se eliminan los datos incompletos, inexactos o irrelevantes. Además, se manejan los valores atípicos y se estandarizan los formatos para asegurar la consistencia y confiabilidad de los datos (Diaz, 2022).

1.2.2 Desarrollar visualizaciones significativas

Selección de Gráficos Adecuados: Es fundamental elegir los tipos de gráficos que mejor representen los datos y faciliten su comprensión. Esto puede incluir gráficos de líneas para mostrar tendencias, gráficos de barras para comparaciones y mapas para visualizar la distribución geográfica de la deforestación (Diaz, 2022).

Organización en Paneles: Al crear un panel de visualización, se agrupan diferentes gráficos y visualizaciones relacionadas en una única interfaz. Esto permite proporcionar una visión global de la deforestación en el Chocó, facilitando la comparación y el análisis de los datos (Diaz, 2022).

1.2.3 Interpretación de resultados

Análisis de Imágenes Visuales: Se analizan las visualizaciones y los patrones identificados para extraer conclusiones significativas sobre las causas y tendencias de la deforestación en el Chocó. Esto puede implicar identificar áreas prioritarias para la conservación, identificar actividades humanas que contribuyen a la deforestación y evaluar el impacto de políticas y programas de conservación (Diaz, 2022).

Identificación de Acciones de Mitigación y Protección: Basándose en los resultados del análisis, se proponen acciones específicas para mitigar la deforestación y proteger los ecosistemas forestales del Chocó. Esto puede incluir la implementación de políticas de conservación, la promoción de prácticas agrícolas sostenibles y la participación comunitaria en la gestión de recursos naturales (Diaz, 2022).

1.2.4 Ciclo de iteración y mejora continua

Revisión y Actualización: Se establece un ciclo de retroalimentación continua para revisar y actualizar los paneles y visualizaciones en función de nueva información y datos proporcionados por el Ministerio de Medio Ambiente y otras fuentes relevantes (Diaz, 2022).

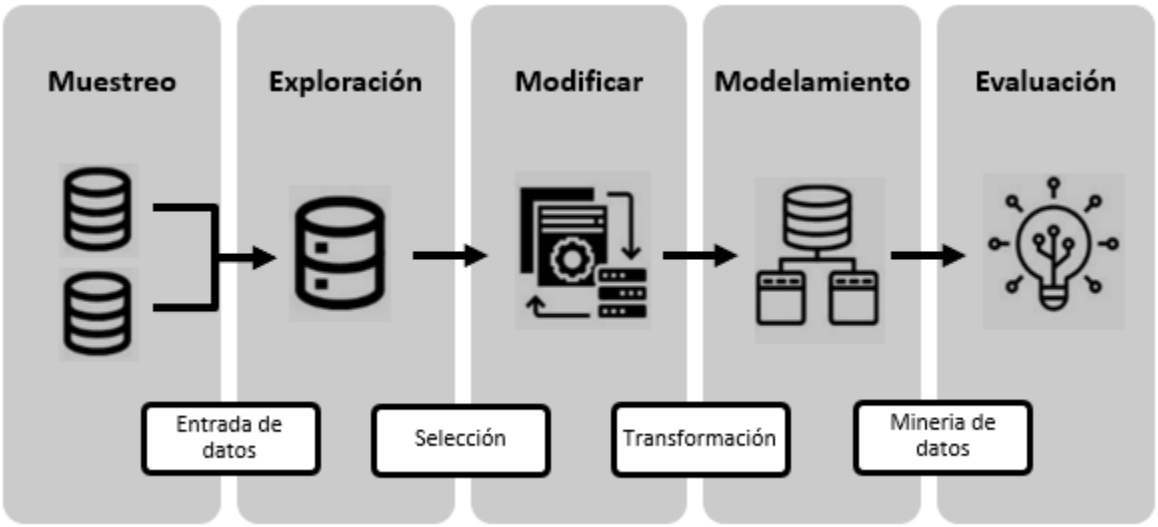
Mejora Iterativa: Se buscan constantemente formas de mejorar la calidad y la utilidad de las visualizaciones y análisis, incorporando retroalimentación de los usuarios y actualizando los métodos y técnicas según sea necesario. Esto garantiza que el proceso de monitoreo y análisis de la deforestación en el Chocó sea dinámico y adaptable a los cambios en el paisaje y las políticas ambientales (Diaz, 2022).

2. Diseño Metodológico

El proyecto ha sido desarrollado con el diseño metodológico SEMMA que permitirá en cada una de sus fases desarrollar e implementar los diferentes procesos utilizados para la creación del tablero de mando que permitirá dar visualizar el análisis sobre las causas de la deforestación en el departamento del Choco. Esta metodología permite organizar estructura y secuencial los procedimientos involucrados, en la obtención de datos, el procesamiento de la información de dichos datos y la creación del diseño que permita exponer las causas relevantes de la deforestación en este departamento (Moine et al., 2011).

Es una metodología de análisis de datos, creada por el SAS Institute, que permite la minería de datos de manera efectiva. Esta metodología se basa en un proceso de selección y modelado de grandes cantidades de datos, en la Figura 4 se definen los pasos que conllevan su desarrollo e implementación lo cual son, Sample (Muestreo), Explore (exploración), Modify (Manipulación), Model (Modela) y Assess (Valoración). Luego se definirá cada paso mencionado (Moine et al., 2011).

Figura 4 Fases Metodología SEMMA



Fuente elaboración propia.

2.1 Fase I - Sample (Muestreo)

Se trata de seleccionar una muestra representativa de los datos que se van a utilizar para la investigación a realizar, por lo cual se realiza la búsqueda y obtención de nuestra principal fuente de información, una base de datos con la información de las causas de la deforestación en el departamento del Choco entre 2014–2021. Esta Base de datos es proporcionada por la página web GOV.CO (AREAS DEFORESTADAS CHOCO | Datos Abiertos Colombia, 2022).

2.2 Fase II - Explore (exploración)

Se refiere a la preparación de los datos que previamente se seleccionaron. En este proceso se validan los datos obtenidos de la base de datos obtenida de la página web gov.co. Identificar cada columna que trae, la clase de datos proporcionados y la cantidad de registros de este. También se validan los datos para identificar anomalías para eliminarlos o transformarlos y poder usarlos en el análisis posterior.

2.3 Fase III - Modify (Manipulación)

En esta fase se realiza la modificación de los datos, en este paso se preparan los datos para realizar el modelado de estos permitiendo así el procesamiento de los datos, teniendo en cuesta esto se validan los datos y eligen los mismos, columnas de la base de datos son relevantes para analizar la deforestación en el choco, que variables se manejarán para identificar la problemática que se quiere mostrar y su estructura.

2.4 Fase IV - Model (Modela)

En esta fase se realiza la aplicación de las técnicas estadístico y aprendizaje automatizado que permite realizar modelos predictivos, siendo esta una fase esencial del proceso de datos en esta metodología, por lo cual se usará la herramienta que permitirá visualizar el análisis de los datos, mostrando así las causas de la deforestación en el departamento del choco. La herramienta que se va a utilizar es Power BI en el que se desarrollara el tablero de control, donde mostrara el diseño en el que se podrán identificar las causas, la cantidad de áreas con deforestación dentro del periodo analizado de los años 2014 a 2021 en el departamento del Choco.

2.5 Fase V – Assess (Valoración)

En esta fase se realiza la evaluación de modelo, realizando la verificación de este, donde evidencie si el modelo dio el resultado esperado de la investigación, es una de las más importantes ya que en ella se realiza la evaluación del proceso realizado, es decir del diseño desarrollado en la herramienta Power BI, se podrá determinar si la implementación de este demostrar lo que se quería identificar y exponer en la visualizar por medio del tablero de control, si la información en el obtenida da solución al objetivo del proyecto desarrollado.

En la tabla 1 se compararán los objetivos expuestos en este proyecto y la metodología SEMMA identificando en cada fase cada objetivo.

Tabla 1 Identificación metodología SEMMA

Objetivo especifico	Fase Metodología SEMMA	Descripción de la actividad
Obtener los datos que me permiten identificar las causas sobre la deforestación en el Chocó durante el período especificado.	Fase I - Sample (Muestreo)	Se plantearon problemáticas ambientales que se tuviera información relevante para realizar un análisis de este. Se identifico que la deforestación es una de las principales. Posteriormente se realizó una validación de varias bases de datos públicas en las que se pudieran identificar las causas de la deforestación en el departamento en el Choco.
	Fase II - Explore (exploración)	Se identifico que la información de la base de datos abiertos de áreas deforestadas del choco, se validaron cada dato, validando que fueran relevantes y limpios, eliminando anomalías y verificando patrones que nos permitieran

		manipularlos. (AREAS DEFORESTADAS CHOCO | Datos Abiertos Colombia, 2022).
Identificar la metodología para el desarrollo de un tablero de Power BI que permita relacionar las variables de deforestación del choco y muestren de manera clara y comprensible la evolución temporal de las causas.	Fase III - Modify (Manipulación)	Se realiza ajustes y transformaciones en los datos de deforestación del Chocó para facilitar su análisis y visualización en el tablero de Power BI
	Fase IV - Model (Modela)	Se realiza la creación del tablero de control en la herramienta Power BI exponiendo sobre este diseño, las casus de la deforestación en el departamento del choco y mostrando en análisis de los mismos datos procesados
Validar las funcionalidades interactivas en el tablero de control de Power BI que permitan a los usuarios identificar las causas de deforestación del Choco de manera dinámica	Fase V – Assess (Valoración)	Se realiza una validación del tablero de control, que en el este expuesta la información requerida para la solución, que este también este diseñado de manera infectiva en la que se puedan ver diferentes puntos y expongan la problemática de manera fácil y entendible.

Fuente: Elaboración propia

3. Resultados

Se desarrolló un plan de acción detallado para recopilar los resultados, el cual se presenta a continuación, comenzando con la obtención de los datos que se extrajeron de la base de datos del gobierno de Colombia sobre las áreas deforestadas en el departamento del Chocó (AREAS DEFORESTADAS CHOCO | Datos Abiertos Colombia, 2022).

La Figura 5 representa la base de datos tomada para elaborar los tableros en power BI mencionados anteriormente, se pueden observar 10 columnas y 7937 filas, en la limpieza de datos se eliminaron 3 columnas las cuales no representaban información relevante para su análisis entre las cuales se encuentra, (tipo de geometría, imagen y observación), estas tres variables al ser revisadas minuciosa mente se concluyeron que su análisis no representaba ningún valor para el presente trabajo. también, se realizó una limpieza exhaustiva de la información para prepararla para la fase de modelado.

Figura 5 Modelado de la información para los tableros de power BI

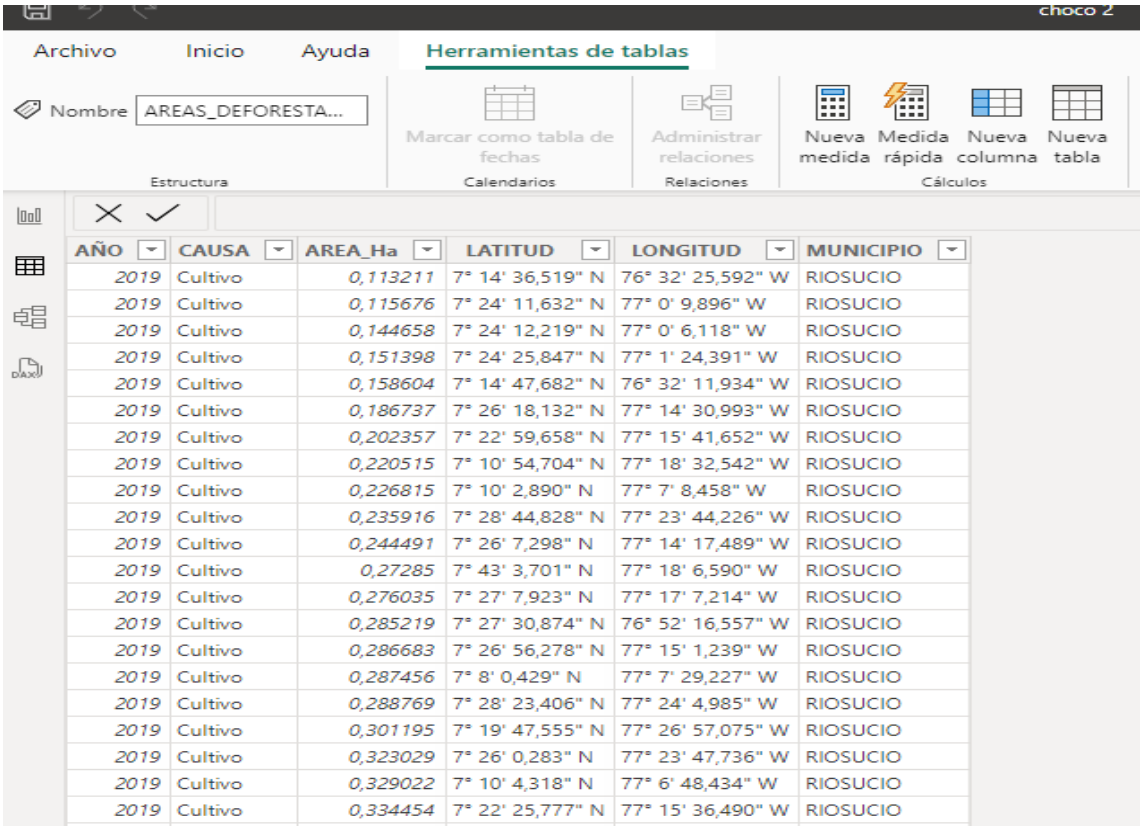
	A	B	C	D	E	F	G	H	I	J
1	ID	TIPO GEOME	AÑO	IMAGEN	CAUSA	AREA_Ha	OBSERVACIO	LATITUD	LONGITUD	MUNICIPIO
2	7463	Polygon	2,02	T18PTQ_202(	Cultivo	0,126512	Ninguna	8° 22' 43,726"	77° 14' 19,019	ACANDÍ
3	4659	Polygon	2,019	T18PTQ_201(	Cultivo	0,14072	Ninguna	8° 30' 47,551"	77° 19' 30,507	ACANDÍ
4	7459	Polygon	2,02	T18PTQ_202(	Cultivo	0,14125	Ninguna	8° 22' 30,966"	77° 13' 49,720	ACANDÍ
5	9719	Polygon	2,021	T18PTQ_202(	Cultivo	0,158066	Ninguna	8° 36' 11.256"	77° 20' 50.757	ACANDÍ
6	7464	Polygon	2,02	T18PTQ_202(	Cultivo	0,184549	Ninguna	8° 22' 22,008"	77° 14' 3,921"	ACANDÍ
7	7445	Polygon	2,02	T18PTQ_202(	Cultivo	0,184854	Ninguna	8° 20' 7,286"	77° 12' 42,123	ACANDÍ
8	7440	Polygon	2,02	T18PTQ_202(	Cultivo	0,185392	Ninguna	8° 23' 18,620"	77° 14' 40,196	ACANDÍ
9	7484	Polygon	2,02	T18PTQ_202(	Cultivo	0,200436	Ninguna	8° 27' 50,102"	77° 21' 38,627	ACANDÍ
10	7478	Polygon	2,02	T18PTQ_202(	Cultivo	0,214739	Ninguna	8° 24' 2,397"	77° 9' 34,027"	ACANDÍ
11	4684	Polygon	2,019	T18PTQ_201(	Cultivo	0,217124	Ninguna	8° 29' 1,740"	77° 22' 8,066"	ACANDÍ
12	4719	Polygon	2,019	T18PTQ_201(	Ganadería	0,234319	Ninguna	8° 28' 33,947"	77° 15' 13,315	ACANDÍ
13	7461	Polygon	2,02	T18PTQ_202(	Cultivo	0,264092	Ninguna	8° 22' 23,660"	77° 14' 3,844"	ACANDÍ
14	7419	Polygon	2,02	T18PTQ_202(	Cultivo	0,269658	Ninguna	8° 31' 23,497"	77° 23' 35,401	ACANDÍ
15	4658	Polygon	2,019	T18PTQ_201(	Cultivo	0,279392	Ninguna	8° 30' 55,153"	77° 19' 39,360	ACANDÍ
16	4726	Polygon	2,019	T18PTQ_201(	Ganadería	0,284863	Ninguna	8° 25' 48,509"	77° 11' 27,234	ACANDÍ
17	7434	Polygon	2,02	T18PTQ_202(	Cultivo	0,287808	Ninguna	8° 28' 48,205"	77° 15' 53,533	ACANDÍ
18	7462	Polygon	2,02	T18PTQ_202(	Cultivo	0,293221	Ninguna	8° 22' 26,779"	77° 14' 7,513"	ACANDÍ
19	4729	Polygon	2,019	T18PTQ_201(	Ganadería	0,295419	Ninguna	8° 26' 28,363"	77° 13' 1,112"	ACANDÍ
20	4739	Polygon	2,019	T18PTQ_201(	Ganadería	0,30872	Ninguna	8° 23' 32,584"	77° 11' 50,315	ACANDÍ
21	4744	Polygon	2,019	T18PTQ_201(	Ganadería	0,313635	Ninguna	8° 23' 15,517"	77° 12' 28,232	ACANDÍ

Fuente elaboración propia

El siguiente paso fue la creación de dos paneles de visualización en Power BI. Fueron diseñados estos paneles con el fin de presentar la información más relevante identificada en el informe propuesto de manera clara y concisa, lo que contribuye a una comprensión más profunda de los datos analizados, con base a la depuración de la base de datos, se procedió con la limpieza, ya que presentaban anomalías en la modelación de los datos, como el nombre de algunas regiones que presidían en diferentes países, fechas que parecían en números decimales, y otras más.

En la figura 6 se evidencia que al momento de modelar los datos, se contaba con 7 columnas de información, entre las cuales se encuentran 2 con las coordenadas de los diferentes municipios, por lo tanto, quedaban 5 columnas, para poder realizar los tableros de power BI con la información relevante para elaborar las gráficas y permitir una mayor comprensión de los datos.

Figura 6 Tabla de Información Power BI



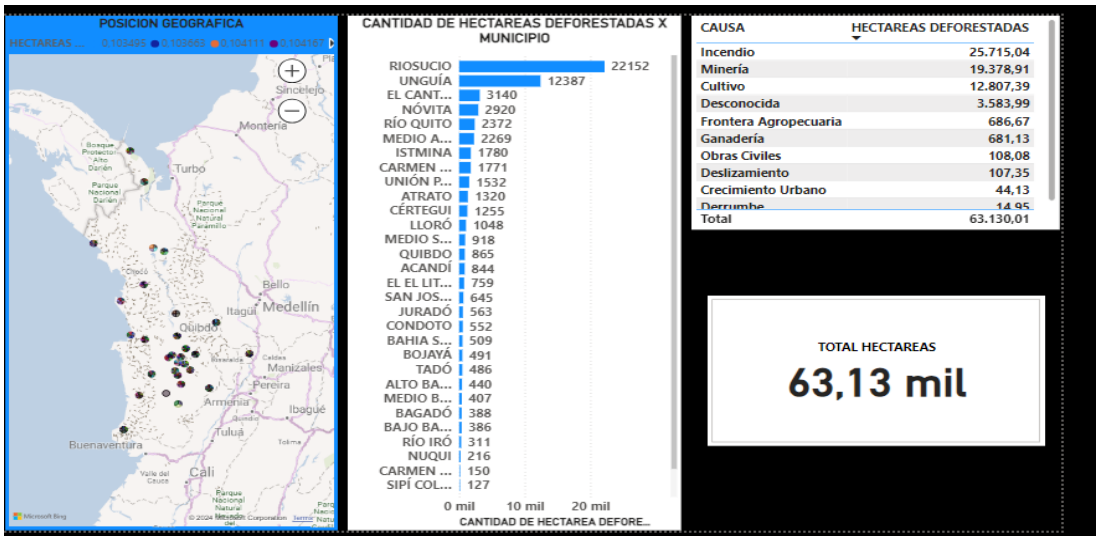
AÑO	CAUSA	AREA_Ha	LATITUD	LONGITUD	MUNICIPIO
2019	Cultivo	0,113211	7° 14' 36,519" N	76° 32' 25,592" W	RIOSUCIO
2019	Cultivo	0,115676	7° 24' 11,632" N	77° 0' 9,896" W	RIOSUCIO
2019	Cultivo	0,144658	7° 24' 12,219" N	77° 0' 6,118" W	RIOSUCIO
2019	Cultivo	0,151398	7° 24' 25,847" N	77° 1' 24,391" W	RIOSUCIO
2019	Cultivo	0,158604	7° 14' 47,682" N	76° 32' 11,934" W	RIOSUCIO
2019	Cultivo	0,186737	7° 26' 18,132" N	77° 14' 30,993" W	RIOSUCIO
2019	Cultivo	0,202357	7° 22' 59,658" N	77° 15' 41,652" W	RIOSUCIO
2019	Cultivo	0,220515	7° 10' 54,704" N	77° 18' 32,542" W	RIOSUCIO
2019	Cultivo	0,226815	7° 10' 2,890" N	77° 7' 8,458" W	RIOSUCIO
2019	Cultivo	0,235916	7° 28' 44,828" N	77° 23' 44,226" W	RIOSUCIO
2019	Cultivo	0,244491	7° 26' 7,298" N	77° 14' 17,489" W	RIOSUCIO
2019	Cultivo	0,27285	7° 43' 3,701" N	77° 18' 6,590" W	RIOSUCIO
2019	Cultivo	0,276035	7° 27' 7,923" N	77° 17' 7,214" W	RIOSUCIO
2019	Cultivo	0,285219	7° 27' 30,874" N	76° 52' 16,557" W	RIOSUCIO
2019	Cultivo	0,286683	7° 26' 56,278" N	77° 15' 1,239" W	RIOSUCIO
2019	Cultivo	0,287456	7° 8' 0,429" N	77° 7' 29,227" W	RIOSUCIO
2019	Cultivo	0,288769	7° 28' 23,406" N	77° 24' 4,985" W	RIOSUCIO
2019	Cultivo	0,301195	7° 19' 47,555" N	77° 26' 57,075" W	RIOSUCIO
2019	Cultivo	0,323029	7° 26' 0,283" N	77° 23' 47,736" W	RIOSUCIO
2019	Cultivo	0,329022	7° 10' 4,318" N	77° 6' 48,434" W	RIOSUCIO
2019	Cultivo	0,334454	7° 22' 25,777" N	77° 15' 36,490" W	RIOSUCIO

Fuente elaboración propia

La Figura 7 representa el primer tablero de visualización, donde se presenta una segmentación de la información por municipio en el departamento del Chocó. En este tablero, se exhiben tanto el número de hectáreas deforestadas por cada municipio como las causas principales asociadas a este fenómeno.

La interacción del tablero permite la ubicación geográfica de cada municipio en el mapa de Colombia. Esta funcionalidad facilita la observación y comprensión de datos específicos, como la cantidad de hectáreas deforestadas y la causa principal de deforestación en cada área municipal. Tal interacción mejora significativamente la claridad y la profundidad del análisis de la información presentada.

Figura 7 Tablero 1 Representación Geográfica



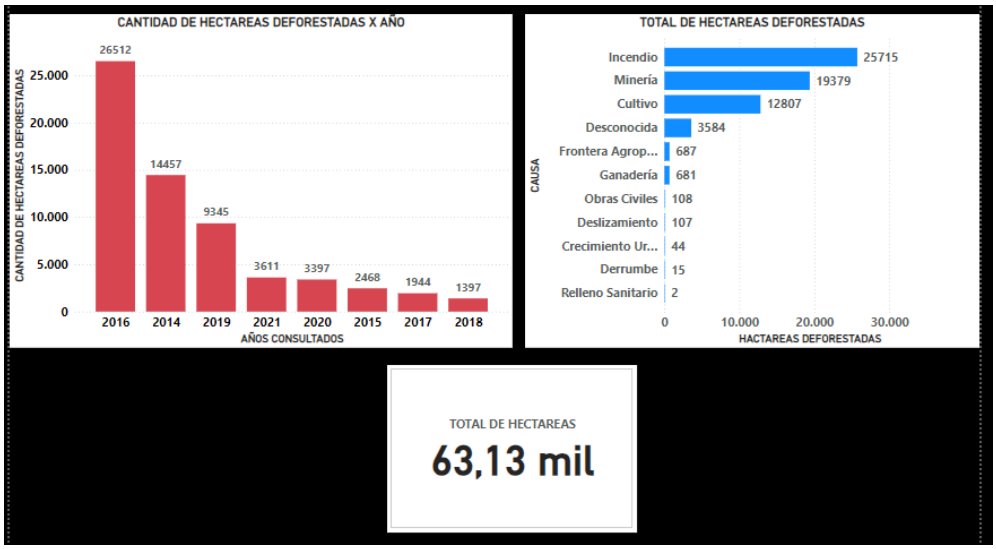
Fuente elaboración propia.

La Figura 8 presenta el segundo tablero de visualización, el cual exhibe la información a través de dos diagramas de barras. Estos gráficos proporcionan una representación visual de las cifras de los datos y su evolución a lo largo de los últimos años.

Basados en el análisis del tablero anterior, se observa un aumento en la deforestación durante los últimos años, siendo la minería la causa principal de este fenómeno con más de 19 mil áreas afectadas y el incendio presentado en el 2016 que afecto un total de 25 mil Ha aprox. Además, se constata un incremento en la cantidad de hectáreas desforestadas durante el período comprendido entre 2014 y 2021.

La interacción del segundo tablero permite identificar la cantidad de hectáreas deforestadas por año y determinar cuál de las causas tuvo un mayor impacto en cada período. Esta funcionalidad proporciona una comprensión más detallada y dinámica del comportamiento de la deforestación a lo largo del tiempo, así como de la influencia de diferentes factores en este proceso.

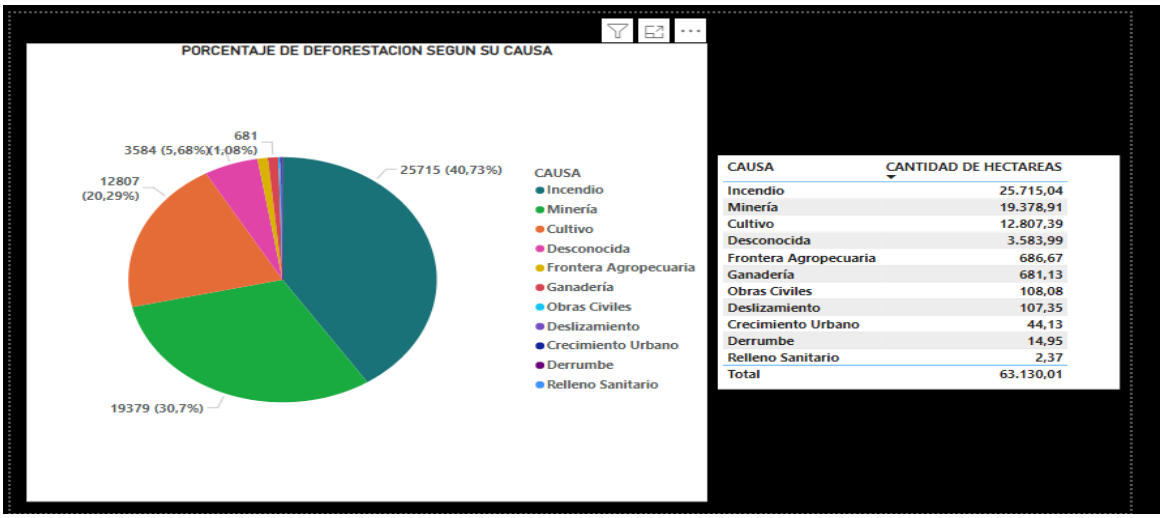
Figura 8 Tablero 2 Afectaciones por año



Fuente elaboración propia.

En la Figura 9, se observa que la causa principal de deforestación en el choco son los incendios con un 40.7% seguido de la minería, con un 30,7% y los cultivos con un 20,3% el restante se segrega en las demás causas, esto muestra que hay tres casos que representan un 91.7% de la deforestación en las cuales los entes gubernamentales deberían tomar cartas en el asunto buscando una solución para mitigar el impacto y los daños que se generan al ambiente.

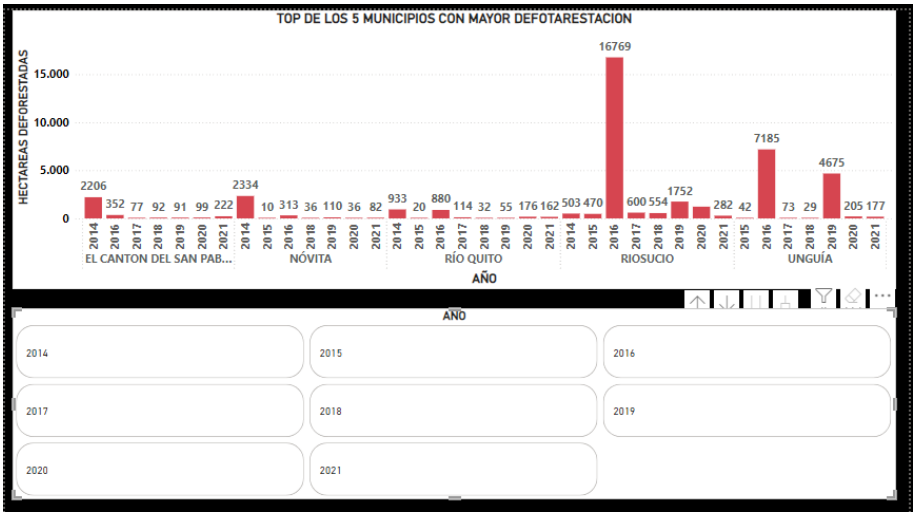
Figura 9 Porcentaje de deforestación según la causa



Fuente elaboración propia.

En la Figura 10 se realizó un filtrado con un top de los 5 municipios con mayor deforestación, adicionalmente se le agrego la opción que permite visualizar como fue el comportamiento del municipio frente a la deforestación en los diferentes años, esto permite saber el comportamiento que ha tenido en cada año, poder determinar si ha disminuido o ha aumentado.

Figura 10 Top 5 de Municipios con mayor deforestación



Fuente elaboración propia.

En base a lo anterior la información modelada en los diferentes tableros creados puede apoyar la toma de decisiones al proporcionar una visualización clara y comprensible de los datos mediante gráficos y tablas, lo que facilita la identificación rápida de tendencias, patrones y anomalías; además, las herramientas de análisis disponibles permiten profundizar en los datos, descubriendo información crucial que podría ser pasada por alto, y el seguimiento del progreso hacia objetivos específicos ayuda a ajustar estrategias en tiempo real. La posibilidad de compartir estos tableros fomenta la comunicación y colaboración entre los usuarios, mientras que el monitoreo del rendimiento de procesos y campañas permite identificar áreas de mejora y optimizar resultados. En conjunto, estas capacidades hacen que los tableros creados en Power BI sean una herramienta versátil que facilita la comprensión de los datos y la toma de decisiones bien informadas.

4. Discusión

El análisis de datos realizado entre 2014 y 2021 en el Departamento del Chocó revela una tendencia alarmante de deforestación, con impactos significativos en el medio ambiente, la biodiversidad y las comunidades locales. Se ha observado un aumento preocupante en la tasa de pérdida de bosques, como lo indica el informe del EDEAM, que muestra una pérdida aproximada de 295,000 hectáreas solo en el año 2021. Esta pérdida de bosques se traduce en una emisión de aproximadamente 159 millones de toneladas de CO2, según datos del Global Forest Watch. En 2019, la deforestación causada por acciones humanas afectó aproximadamente 177,644 hectáreas.

Este análisis se centra en comprender las causas de la deforestación, que van desde la expansión de la agricultura y la ganadería hasta la minería ilegal y la deforestación a gran escala. Se busca determinar estas causas y evaluar el alcance de la deforestación en las áreas afectadas y relacionadas, sus impactos en la biodiversidad y en las comunidades indígenas y afroamericanas, cuyas vidas y culturas están ligadas a los recursos naturales.

En consecuencia, este análisis de datos ofrece una perspectiva detallada sobre la deforestación en el Chocó, destacando la importancia de abordar este problema de manera integral y urgente para proteger los valiosos recursos naturales y culturales de la región.

5.Conclusiones y recomendaciones

5.1 Conclusiones

El desarrollo de paneles de control Power BI permite una visualización efectiva de las causas y evolución en el tiempo de la deforestación en el Chocó, por lo cual la interactividad y transparencia de la presentación de datos facilita la comprensión y el análisis, apoyando así la toma de decisiones informadas para abordar y mitigar las causas de la deforestación en la región, permitiendo poder implementar este método ayudando a identificar y analizar de forma clara y sencilla las causas de la deforestación en el Chocó durante el periodo 2014-2021.

El análisis de los datos arroja evidencia de que los tres municipios con mayor influencia y/o impacto son el Riosucio, Unguía y el Canton de San Pablo., también descubrió acres de deforestación y deforestación en curso debido a la tala de vegetación para abrir minas y construir infraestructura relacionada, como carreteras y campamentos mineros.

La integración de funciones interactivas en los paneles de Power BI brindará a los usuarios la capacidad de explorar datos dinámicamente, ayudándolos a tomar decisiones informadas sobre la protección y gestión de los bosques en su región, según los resultados acordados entre 2014 y 2021, se determinó que la principal causa de deforestación en el Chocó son los incendios (40.7%) y la minería ilegal (20.3%).

Las pruebas de la funcionalidad interactiva del panel de Power BI confirmaron que las causas de la deforestación en Chocó se pueden determinar dinámicamente. La capacidad de filtrar y segmentar datos de manera efectiva, así como visualizaciones claras y detalladas, permite una comprensión profunda y un análisis detallado de los factores que impulsan la deforestación por lo cual este tablero se presenta como una valiosa herramienta para la toma de decisiones e implementación de medidas ambientales en la región.

5.2 Recomendaciones

La investigación propone un enfoque basado en el análisis de datos para abordar el problema de la deforestación en el Chocó. Sugiere abandonar los supuestos cualitativos en favor de explicaciones objetivas y cuantitativas sobre el problema. Se identifica que la deforestación está arraigada en bases sociales y culturales relacionadas con la minería ilegal y el cultivo de coca.

Se destaca la importancia de abordar este problema desde la perspectiva de la "seguridad humana", considerando la seguridad en un contexto multilateral. Se plantean preguntas para futuras investigaciones, como las variaciones en los comportamientos criminales en diferentes áreas boscosas y los subtipos que influyen en estas variaciones.

Se enfatiza que las respuestas a estas preguntas deben orientar hacia el desarrollo de metas y estrategias para las políticas de defensa y seguridad, responsabilidad tanto del estado como de la humanidad. Se advierte que la deforestación sistemática tiene consecuencias graves, como la violación de los derechos humanos a un ambiente saludable, la protección de los recursos naturales y la preservación del medio ambiente como un recurso único.

Bibliografía

- Forero, D. E. (2023). Deforestación en el departamento del chocó: causas, datos y enfoques asociados al concepto de seguridad multidimensional. Recuperado de: http://hdl.handle.net/10654/45311.
- Restrepo, J. [Jhon David restrepo Berrío]. (2022). Evaluación de los factores asociados con la deforestación en los municipios de Riosucio y río quito, chocó, Colombia. Evaluación de los factores asociados con la deforestación en los municipios de Riosucio y río quito, chocó, Colombia. Recuperado 3 de mayo de 2024, de https://manglar.uninorte.edu.co/bitstream/handle/10584/11445/1002089520.pdf?sequence=1&isAllowed=y
- El chocó biogeográfico de Colombia - colección ecológica del banco de occidente. (S. F.). https://www.imeditores.com/banocc/choco/cap3.htm#:~:text=Los%20manglares%20del%20Choc%C3%B3%20Biogeogr%C3%A1fico,de%20manglares%20en%20el%20pa%C3%ADs
- Kanninen, M. [Markku Kanninen], Murdiyarso, D. [Daniel Murdiyarso], Seymour, F. [Frances Seymour], Angelsen, A. [Arild Angelsen], Wunder, S. [Sven Wunder], & German, L. [Laura German]. (2086). ¿Crecen los árboles sobre el dinero?: Implicaciones de la investigación sobre deforestación en las medidas para promover la REDD. Center for International Forestry Research. https://www.cifor-icraf.org/publications/pdf_files/Books/BKanninen0801SP.pdf
- National geographic. (2024, 20 marzo). El último medio siglo ha supuesto el mayor coste de deforestación de la historia de la humanidad, arrasando un 15 por ciento de la superficie mundial de vegetación, equivalente al territorio de España, Portugal y Francia. Https://www.Nationalgeographic.Es/medio-ambiente/deforestacion
- Castañeda, G. [Grecia-Castañeda rojas], canales, S. [Schleiffer-canales Carreño], Ovalle, C. [Christian Ovalle], rabanal, E. [Erick Humberto rabanal Chávez], & universidad privada del norte,peru [universidad privada del norte,peru]. (2023). Intelligent predictive model applying data mining strategies for a credit evaluation of a commercial company. Intelligent predictive model applying data mining strategies

for a credit evaluation of a commercial company. Recuperado 3 de mayo de 2024, de https://laccei.org/LACCEI2023-BuenosAires/papers/Contribution_1148_a.pdf

- Moine, J. M., Haedo, A. S., & Gordillo, S. E. (2011). Estudio comparativo de metodologías para minería de datos. https://sedici.unlp.edu.ar/handle/10915/20034
- Natalia, H. B., & Camilo, R. L. J. (2022). Evaluación de los factores asociados con la deforestación en los municipios de Riosucio y Río Quito, Chocó, Colombia. Universidad del Norte. http://hdl.handle.net/10584/11445
- Davidiseminger. (2024, 22 marzo). ¿Qué es Power BI? - Power BI. Microsoft Learn. https://learn.microsoft.com/es-es/power-bi/fundamentals/power-bi-overview
- AREAS DEFORESTADAS CHOCO | Datos abiertos Colombia. (2022, 20 mayo). https://www.datos.gov.co/Ambiente-y-Desarrollo-Sostenible/AREAS-DEFORESTADAS-CHOCO/iczg-dyt3/data_preview
- Díaz Vásquez, R. A., Acosta Espinoza, J. L., & Checa Cabrera, M. A. (2022). Power bi como herramienta de apoyo a la toma de decisiones. Universidad Y Sociedad, 14(S3), 195–207. Recuperado a partir de https://rus.ucf.edu.cu/index.php/rus/article/view/2949
- Asprilla Panesso, E. S. ., Lozano Urrego, H. Y. ., Bechara Palacio, R. ., & Emiro Ledezma, R. . (2019). El medio ambiente como víctima del conflicto armado en el departamento del Choco?. Pensamiento Americano, 12(23). https://doi.org/10.21803/pensam.v12i22.251